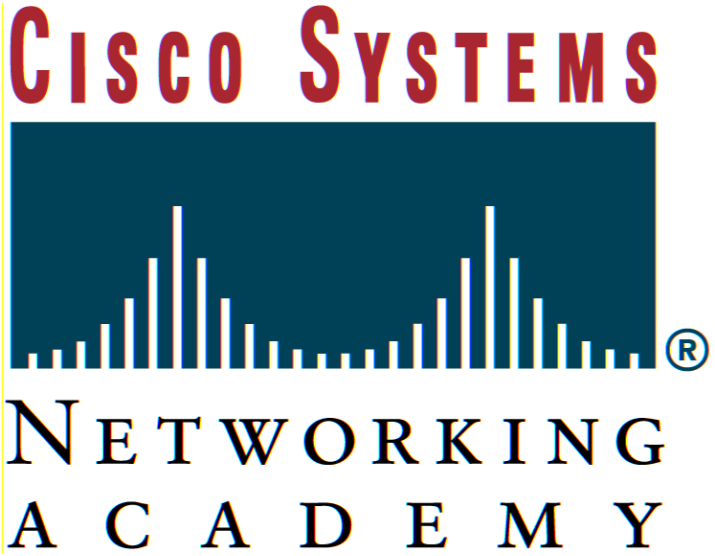
Cabrillo College



Basic BGP Lab Examples

Rick Graziani, Instructor

November 15, 2001

Note to instructors

- If you have downloaded this presentation from the Cisco Networking Academy Community FTP Center, this may not be my latest version of this PowerPoint.
- For the latest PowerPoints for all my CCNA, CCNP, and Wireless classes, please go to my web site:

  http://www.cabrillo.edu/~rgraziani/

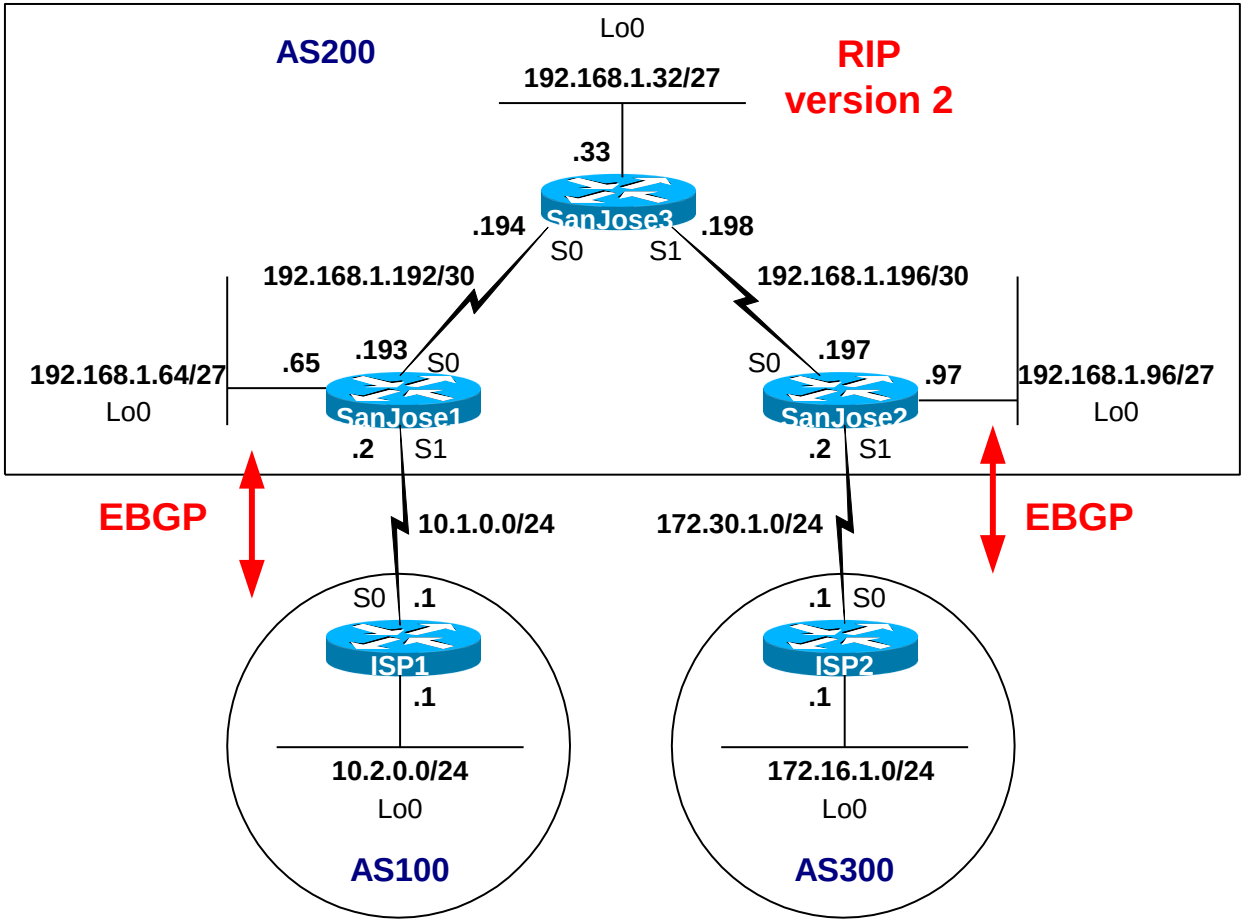
  - The username is *cisco* and the password is *perlman* for all of my materials.
- If you have any questions on any of my materials or the curriculum, please feel free to email me at graziani@cabrillo.edu (I really don't mind helping.) Also, if you run across any typos or errors in my presentations, please let me know.
- I will add "(Updated – *date*)" next to each presentation on my web site that has been updated since these have been uploaded to the FTP center.

Thanks! Rick

Objective

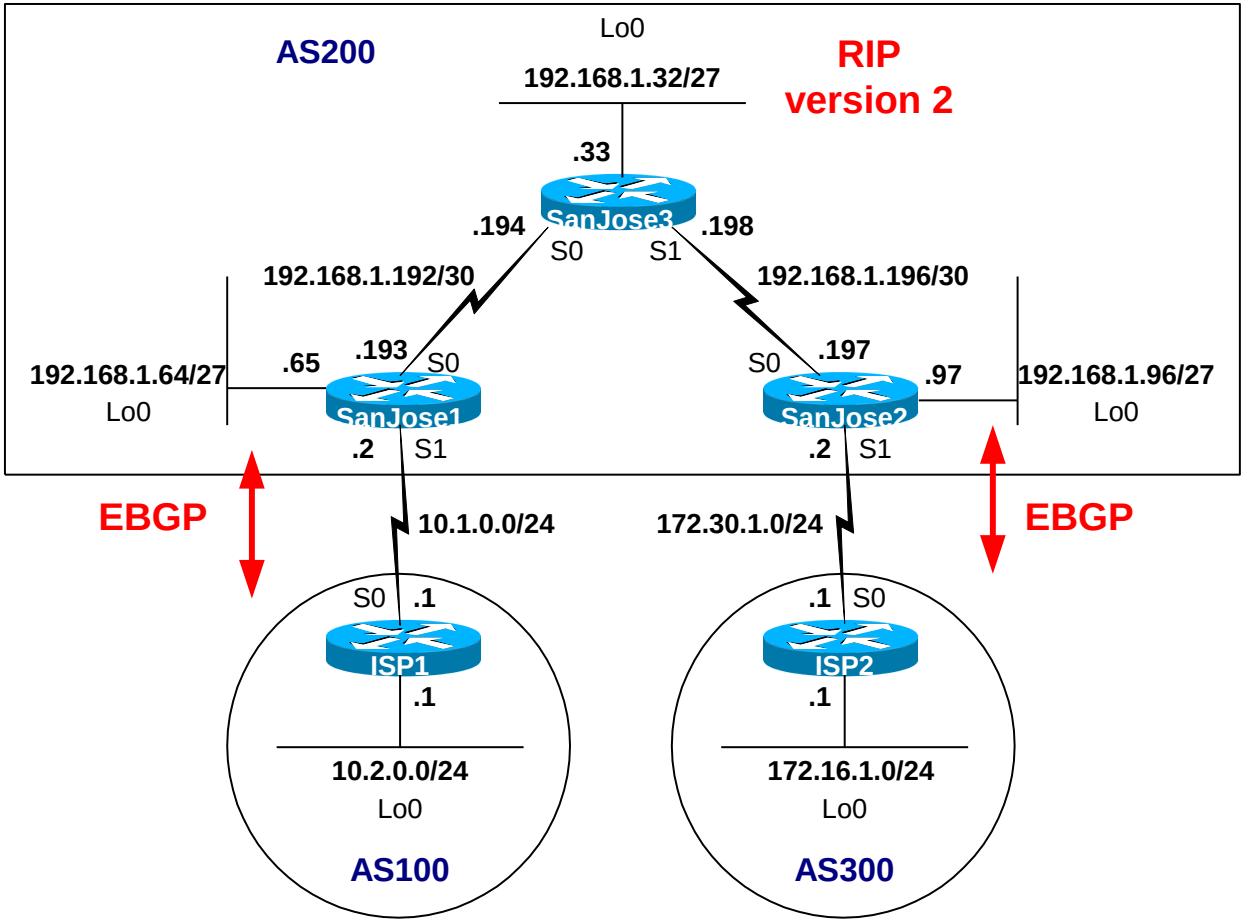
- These lab examples are not meant to demonstrate proper or realistic BGP configurations, nor does it include all the necessary commands to ensure proper routing operation.
- The objective of these examples are to show some of the basic BGP configuration commands and their affect on a network, including wrong configurations or misconfigurations.
- Hopefully this series of configurations and outputs will help explain the how BGP is implemented and its operation.

Our Scenario



Why IBGP?

"With very few exceptions, interior BGP (IBGP) – BGP between peers in the same AS – is used only in multihomed scenarios." – Doyle

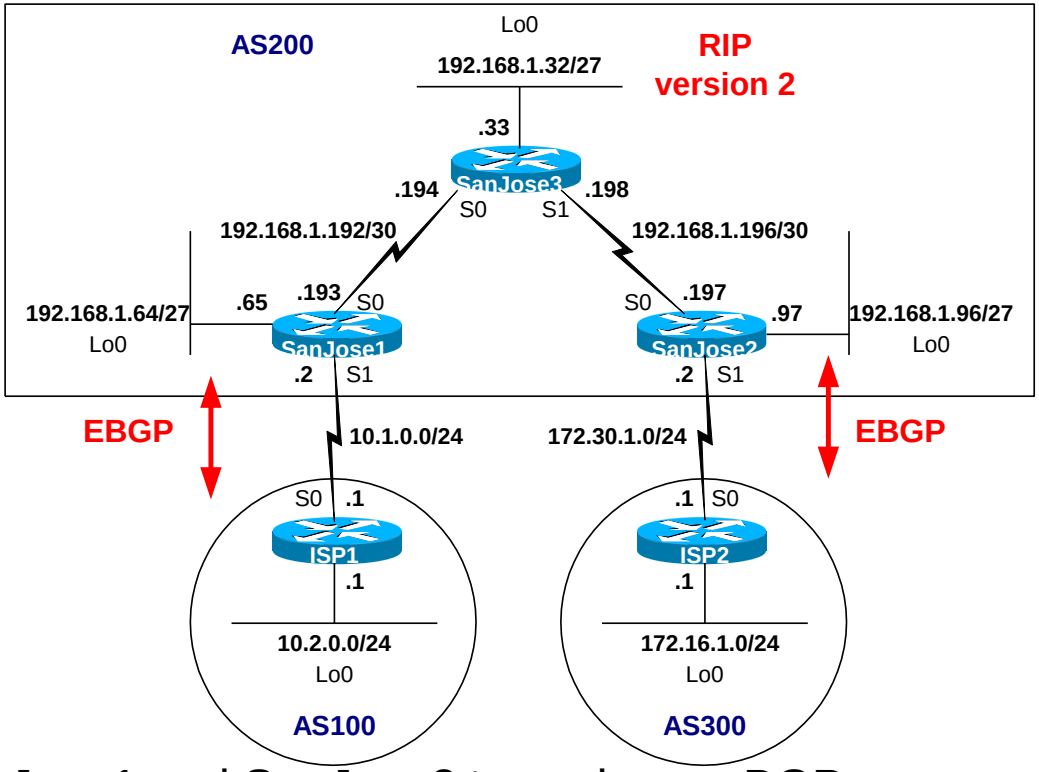


In order for the boundary routers SanJose1 and SanJose2 to exchange BGP routes that they learn from ISP1 and ISP2 respectively, they must do one of the following:

- **Redistribution – Not recommended!**
  - Redistribute from BGP into the IGP (RIPv2) so those routes get propagated to the other boundary router.
  - Redistribute from IGP (RIPv2) to BGP to learn routes from the other boundary router or the use of BGP network statements.
  - **Disadvantages**: Can be a lot of routes being injected into the IGP; may also require the use of route filters so not all routes are injected into the IGP or back out to the other ISP.

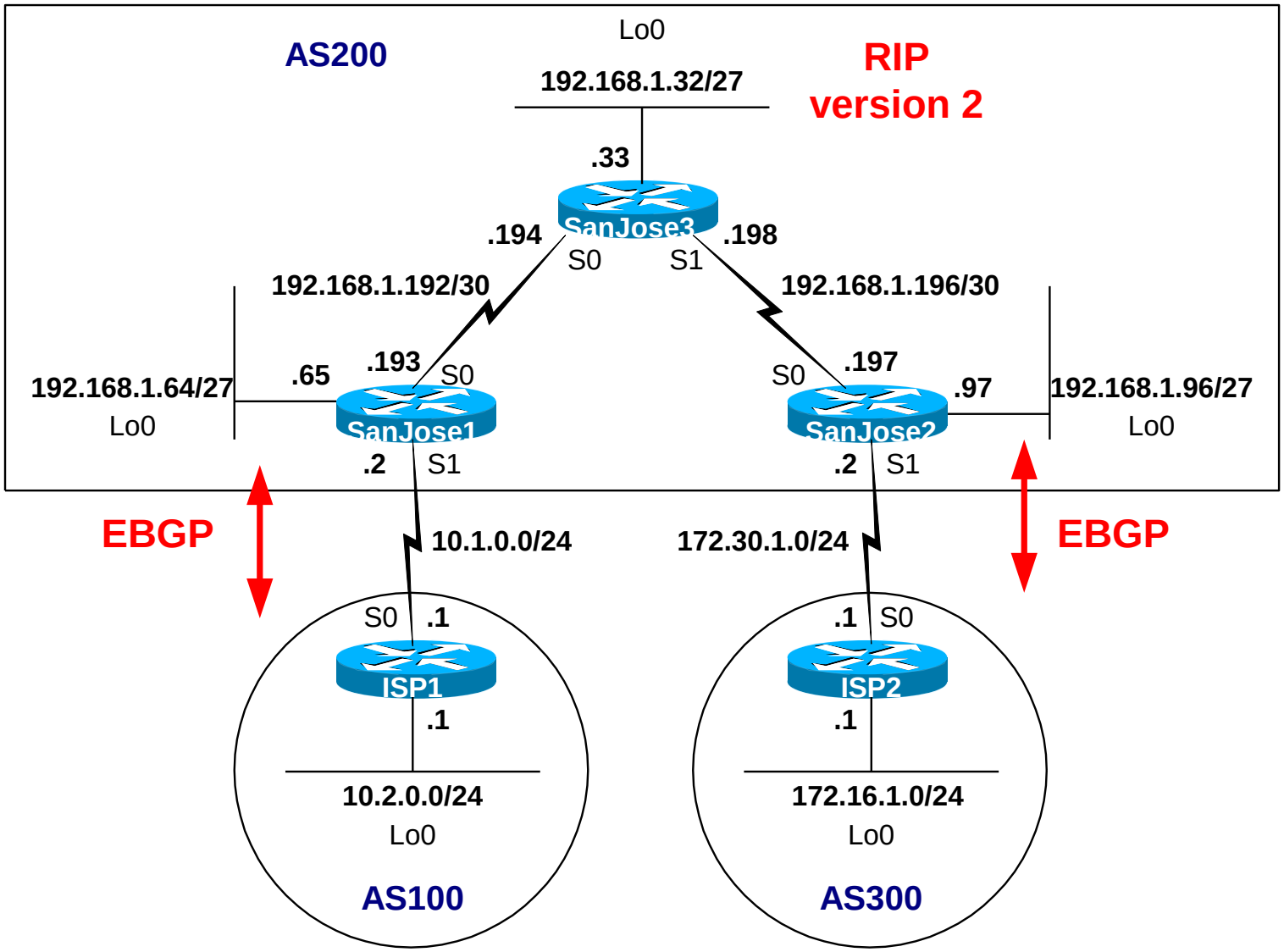
Why IBGP?

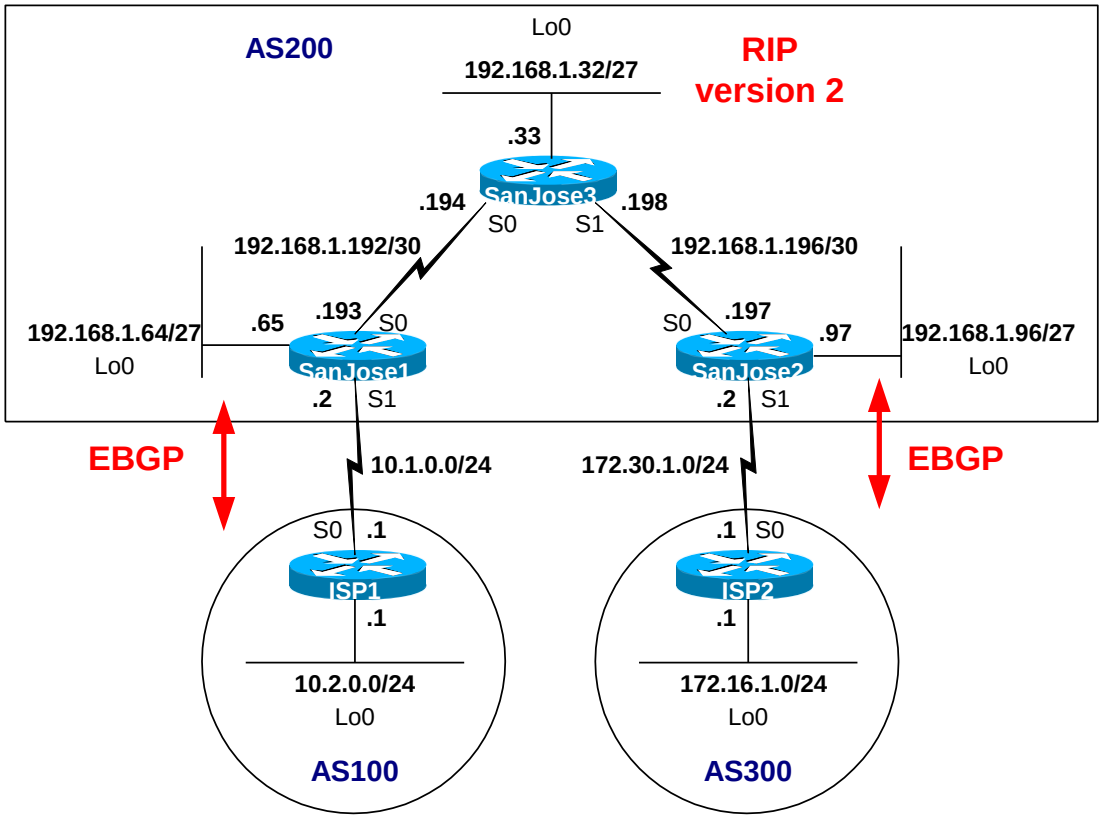
"With very few exceptions, interior BGP (IBGP) – BGP between peers in the same AS – is used only in multihomed scenarios." – Doyle



In order for the boundary routers SanJose1 and SanJose2 to exchange BGP routes that they learn from ISP1 and ISP2 respectively, they must do one of the following:

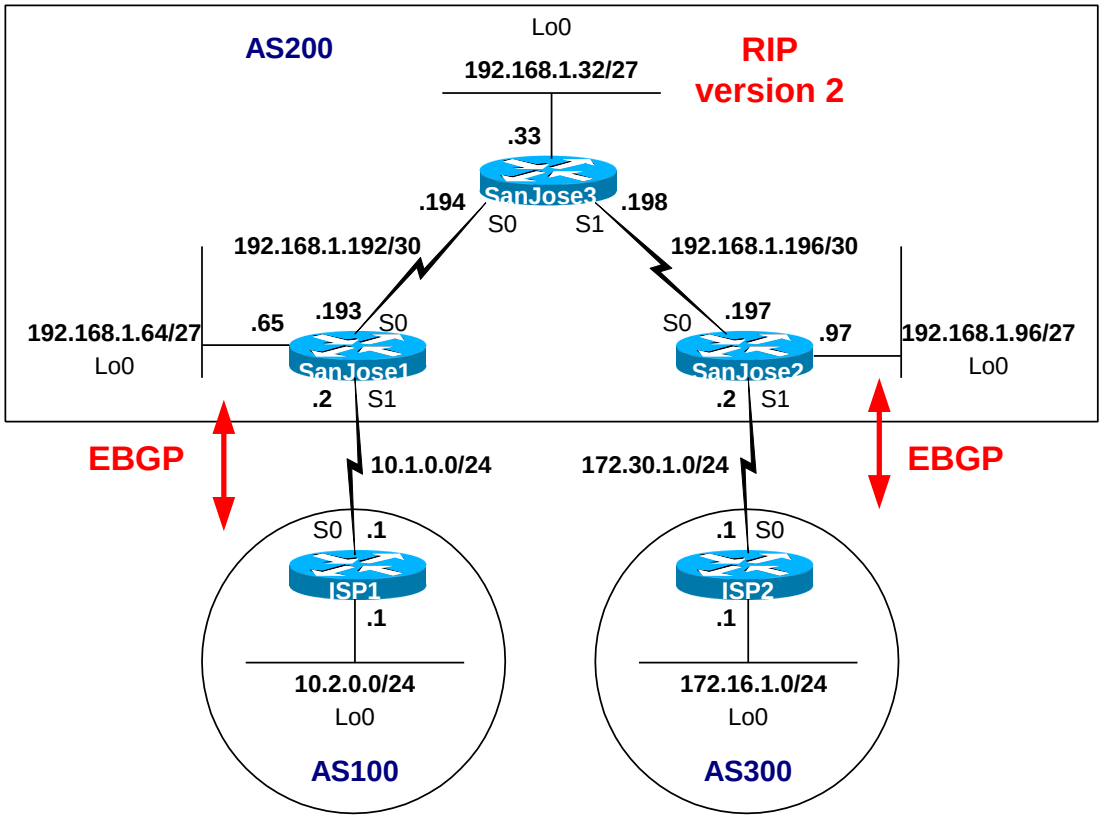
- **IBGP**
  - This allows the boundary routers to share EBGP routes without the redistributing via an IGP.
  - Requires:
    - **Synchronization**: Can't advertise via EBGP if the boundary router doesn't have it in its IGP routing table because the other IGP routers will drop it. (Can turn this off if fully meshed IBGP or non-transit AS!)
    - Must know how to reach the next hop address of the route. (example coming!)


AS200
Lo0
192.168.1.32/27
RIP
version 2
.33
SanJose3
.194
S0
S1
.198
192.168.1.192/30
192.168.1.196/30
192.168.1.64/27
Lo0
.65
.193
S0
SanJose1
.2
S1
S0
.197
SanJose2
.2
S1
.97
192.168.1.96/27
Lo0
EBGP
EBGP
10.1.0.0/24
172.30.1.0/24
S0
.1
ISP1
.1
10.2.0.0/24
Lo0
AS100
.1
S0
ISP2
.1
172.16.1.0/24
Lo0
AS300



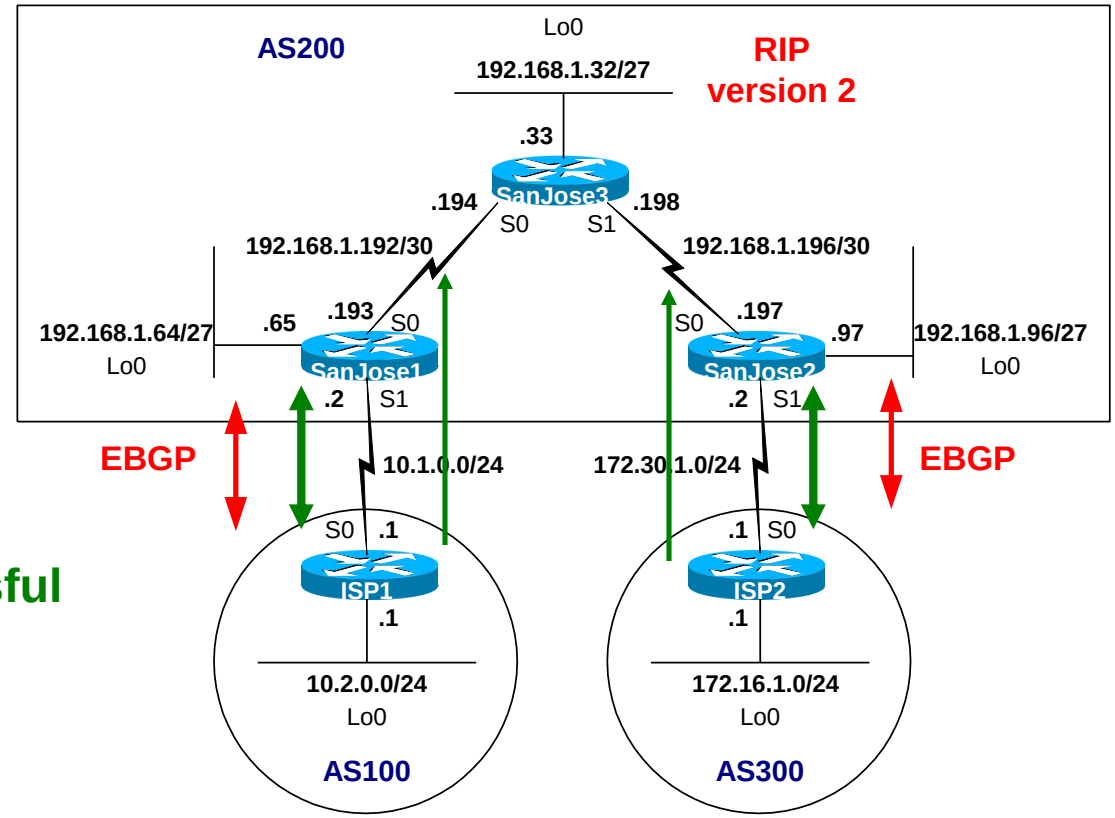
Setup

- RIPv2 running within AS200
- 10.1.0.0/24 is between SanJose1 and ISP1's 10.2.0.0/24 networks, both of which are part of 10.0.0.0/8 classful network.
- RIPv2 was chosen because this will create a different situation than with the two different Class B networks between SanJose2 and ISP2.



Scenario A: Step #1 – EBGP and IGP (RIPv2)

- EBGP is configured between ISP routers and AS200 boundary routers.
- RIPv2 is running within AS200
- RIP network statements only include those networks in AS200
- **No** IBGP has been configured.
- There are **no** static routes.



Results:

- AS 200 Boundary (SanJose1 and SanJose2) routers can ping peer ISP networks
- AS 200 Boundary (SanJose1 and SanJose2) routers can NOT ping opposite ISP networks (SanJose1 can not ping 172.16.1.1)
- ISP routers (ISP1 and ISP2) can ping remote peer's attached networks only, NOT SanJose3 network (we will see why in a moment).

ISP1

```
router bgp 100
 network 10.2.0.0 mask 255.255.255.0
 neighbor 10.1.0.2 remote-as 200
```

ISP2

```
router bgp 300
 network 172.16.0.0
 neighbor 172.30.1.2 remote-as 200
```

SanJose3

```
router rip
 version 2
 network 192.168.1.0
```

SanJose1

```
router rip
 version 2
 network 192.168.1.0
!
router bgp 200
 network 192.168.1.0
 neighbor 10.1.0.1 remote-as 100
```

SanJose2

```
router rip
 version 2
 network 192.168.1.0
!
router bgp 200
 network 192.168.1.0
 neighbor 172.30.1.1 remote-as 300
```

```
ISP1#show ip route

     10.0.0.0/24 is subnetted, 2 subnets
C       10.2.0.0 is directly connected, Loopback0
C       10.1.0.0 is directly connected, Serial0
B    192.168.1.0/24 [20/0] via 10.1.0.2, 00:25:16

ISP1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      0.0.0.0                  0         32768 i
*> 192.168.1.0      10.1.0.2                 0             0 200 i
------------------------------------------------------------------------
ISP2#show ip route

     172.16.0.0/24 is subnetted, 1 subnets
C       172.16.1.0 is directly connected, Loopback0
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/0
B    192.168.1.0/24 [20/0] via 172.30.1.2, 00:30:51

ISP2#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 172.16.0.0       0.0.0.0                  0         32768 i
*> 192.168.1.0      172.30.1.2               0             0 200 i
```

ISPs only see the BGP routes distributed by themselves and their peers in their BGP tables.

SanJose1 only sees the BGP routes distributed by itself and its peer in their BGP tables.

```
SanJose1#show ip route

     10.0.0.0/24 is subnetted, 2 subnets
B       10.2.0.0 [20/0] via 10.1.0.1, 00:32:26
C       10.1.0.0 is directly connected, Serial0/1
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/2] via 192.168.1.194, 00:00:11, Serial0/0
C       192.168.1.64/27 is directly connected, Loopback0
R       192.168.1.32/27 [120/1] via 192.168.1.194, 00:00:11, Serial0/0
C       192.168.1.192/30 is directly connected, Serial0/0
R       192.168.1.196/30 [120/1] via 192.168.1.194, 00:00:11, Serial0/0

SanJose1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      10.1.0.1                 0             0 100 i
*> 192.168.1.0      0.0.0.0                  0         32768 i
```

Status codes - Status of the table entry. The status is displayed at the beginning of each line in the table. It can be one of the following values:

- s —The table entry is suppressed.
- * —The table entry is valid.
- > —The table entry is the best entry to use for that network.
- i —The table entry was learned via an internal BGP (iBGP) session

```
SanJose2#show ip route

B    172.16.0.0/16 [20/0] via 172.30.1.1, 00:27:08
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/1
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
C       192.168.1.96/27 is directly connected, Loopback0
R       192.168.1.64/27 [120/2] via 192.168.1.198, 00:00:09, Serial0/0
R       192.168.1.32/27 [120/1] via 192.168.1.198, 00:00:09, Serial0/0
R       192.168.1.192/30 [120/1] via 192.168.1.198, 00:00:09, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/0

SanJose2#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 172.16.0.0       172.30.1.1               0             0 300 i
*> 192.168.1.0      0.0.0.0                  0         32768 i
```

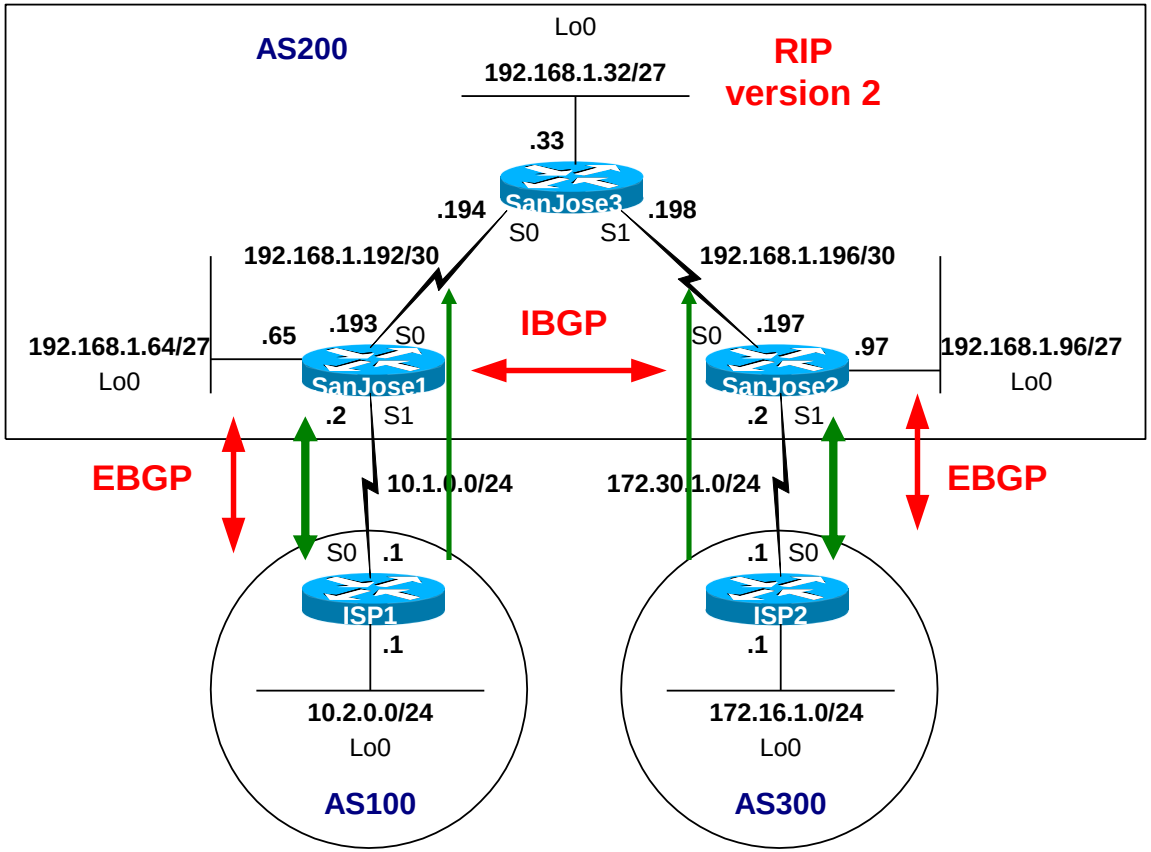
SanJose2 only sees the BGP routes distributed by itself and its peer in their BGP tables.

```
SanJose3#show ip route

     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/1] via 192.168.1.197, 00:00:18, Serial0/1
R       192.168.1.64/27 [120/1] via 192.168.1.193, 00:00:07, Serial0/0
C       192.168.1.32/27 is directly connected, Loopback0
C       192.168.1.192/30 is directly connected, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/1
```

SanJose3 does not have a route to the ISP networks or the networks between the ISPs and the Boundary routers, so the ICMP Echo Replies fail.

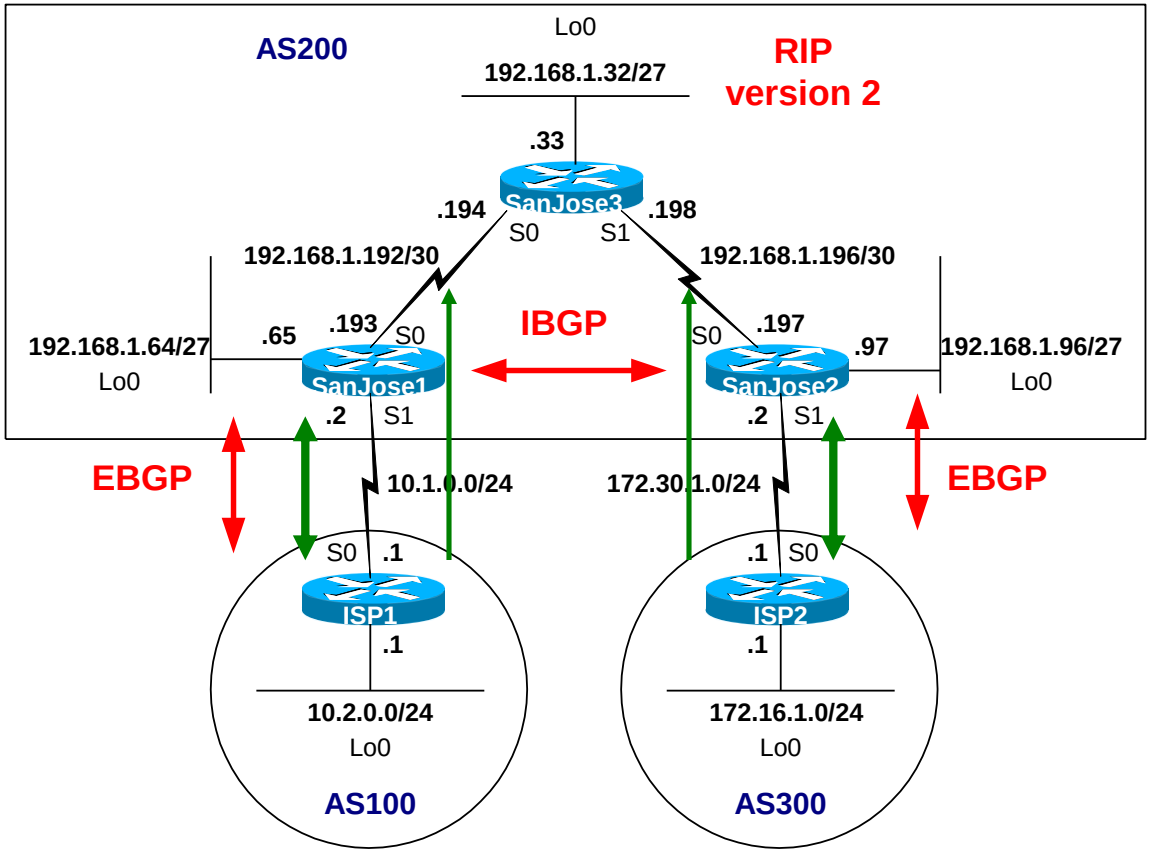
SanJose3 does not get any routes from the boundary routers except IGP routes.

There are no BGP routes redistributed into the IGP and SanJose3 is not IBGP peering with any BGP speakers.



Step #2 – EBGP, IBGP and IGP (RIPv2)

- EBGP is configured between ISP routers and AS200 boundary routers.
- RIPv2 is running within AS200
- IBGP is configured between SanJose1 and SanJose3.
- There are **no** static routes.



Results:

- AS 200 Boundary routers can ping peer ISP networks
- AS 200 Boundary routers can NOT ping opposite ISP networks
- ISP routers can ping remote peer's attached networks only, NOT SanJose3 network
- Synchronization is on, "The BGP synchronization rule states that if an AS provides transit service to another AS, BGP should not advertise a route until all of the routers within the AS have learned about the route via an IGP." The boundary routers do not forward the "opposite" ISP routes because the other routers in the AS do not know about it.

ISP1

```
router bgp 100
 network 10.2.0.0 mask 255.255.255.0
 neighbor 10.1.0.2 remote-as 200
```

ISP2

```
router bgp 300
 network 172.16.0.0
 neighbor 172.30.1.2 remote-as 200
```

SanJose3

```
router rip
 version 2
 network 192.168.1.0
```

SanJose1

```
router rip
 version 2
 network 192.168.1.0
!
router bgp 200
 network 192.168.1.0
 neighbor 10.1.0.1 remote-as 100
 neighbor 192.168.1.197 remote-as 200
```

SanJose2

```
router rip
 version 2
 network 192.168.1.0
!
router bgp 200
 network 192.168.1.0
 neighbor 172.30.1.1 remote-as 300
 neighbor 192.168.1.193 remote-as 200
```

SanJose1 still only sees the BGP routes distributed by itself and its peer in their BGP tables.

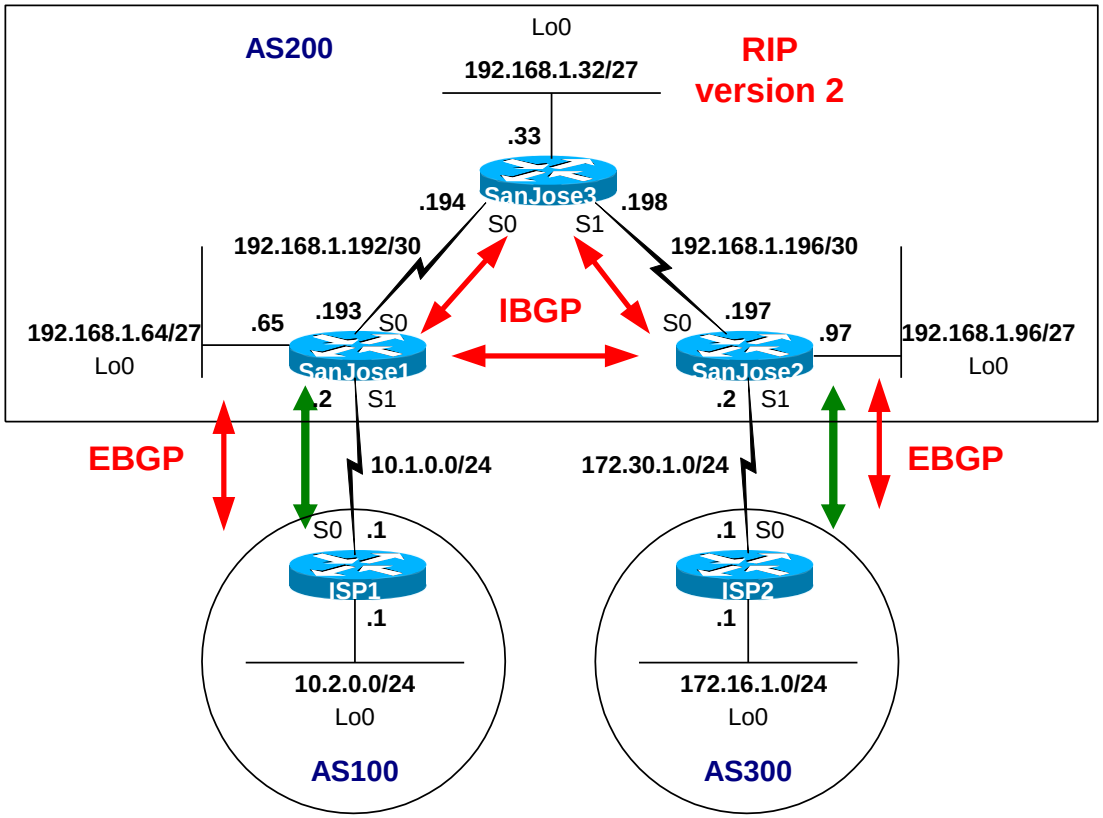
```
SanJose1#show ip route

10.0.0.0/24 is subnetted, 2 subnets
B       10.2.0.0 [20/0] via 10.1.0.1, 00:09:13
C       10.1.0.0 is directly connected, Serial0/1
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/2] via 192.168.1.194, 00:00:27, Serial0/0
C       192.168.1.64/27 is directly connected, Loopback0
R       192.168.1.32/27 [120/1] via 192.168.1.194, 00:00:27, Serial0/0
C       192.168.1.192/30 is directly connected, Serial0/0
R       192.168.1.196/30 [120/1] via 192.168.1.194, 00:00:27, Serial0/0

SanJose1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      10.1.0.1                 0             0 100 i
* i172.16.0.0       172.30.1.1               0    100      0 300 i
*> 192.168.1.0      0.0.0.0                  0         32768 i
* i                 192.168.1.197            0    100      0 i
```

- **Even though 172.16.0.0 network is in BGP table it is not in routing table because there is no route to the next hop of 172.30.1.1**
- **Also, the BGP Synchronization rule keeps the BGP route from being in the routing table, since there is no IGP route.**

```
SanJose2#show ip route

B     172.16.0.0/16 [20/0] via 172.30.1.1, 00:11:27
      172.30.0.0/24 is subnetted, 1 subnets
C        172.30.1.0 is directly connected, Serial0/1
      192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
C        192.168.1.96/27 is directly connected, Loopback0
R        192.168.1.64/27 [120/2] via 192.168.1.198, 00:00:19, Serial0/0
R        192.168.1.32/27 [120/1] via 192.168.1.198, 00:00:19, Serial0/0
R        192.168.1.192/30 [120/1] via 192.168.1.198, 00:00:19, Serial0/0
C        192.168.1.196/30 is directly connected, Serial0/0

SanJose2#show ip bgp
Network          Next Hop            Metric LocPrf Weight Path
* i10.2.0.0/24      10.1.0.1                 0    100      0 100 i
*> 172.16.0.0       172.30.1.1               0             0 300 i
* i192.168.1.0      192.168.1.193            0    100      0 i
*>                  0.0.0.0                  0         32768 i
```

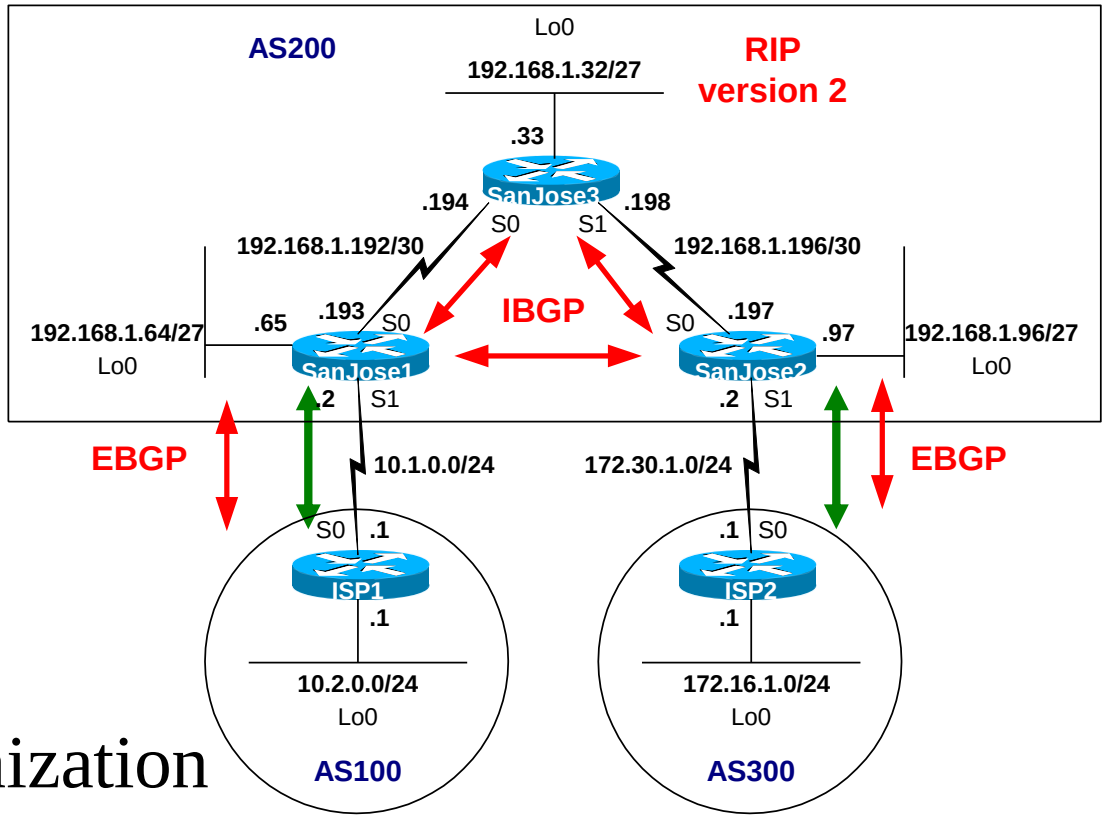
SanJose2 still only sees the BGP routes distributed by itself and its peer in their BGP tables.

- **Even though 10.2.0.0 network is in BGP table it is not in routing table because there is no route to the next hop of 10.1.0.1**
- **Also, the BGP Synchronization rule keeps the BGP route from being in the routing table, since there is no IGP route.**



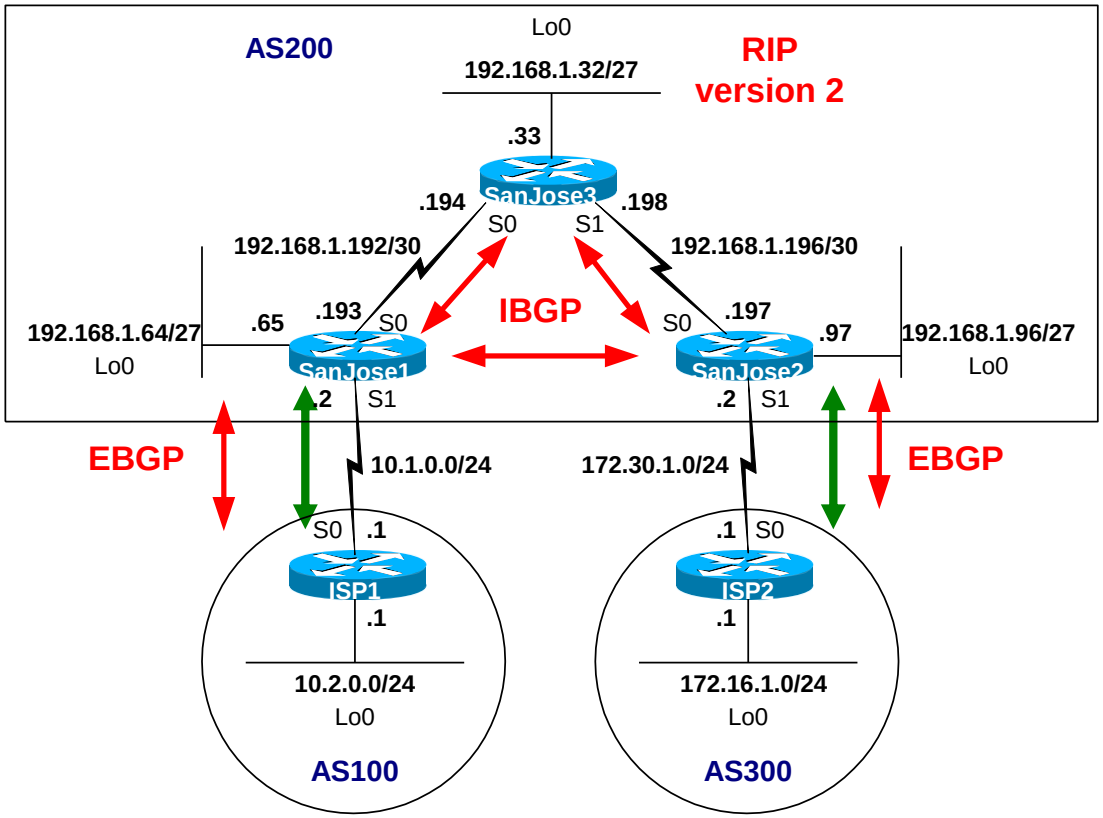
Step #3 – Adding SanJose3 as a IBGP peer

- EBGP is configured between ISP routers and AS200 boundary routers.
- RIPv2 is running within AS200
- IBGP is configured between SanJose1, SanJose2 and SanJose3, full-mesh.
- No synchronization has been configured on all IBGP routers, because we are fully meshed IBGP.
- There are **no** static routes.



Synchronization

- When an IBGP router receives an update about a destination from an IBGP peer, it tries to verify reachability to that destination via an **IGP**, such as RIP or OSPF.
- If the IBGP router can't find the destination network in it's **IGP** routing table, it **will not** advertise the destination to other BGP peers.
- If the route is **not** reachable through the **IGP** running within the AS, non-BGP routers will not be able to route traffic passing through the AS towards this destination.
- It is pointless to advertise destinations to external peers if traffic sent through this AS is going to be dropped by some non-BGP router within the AS anyway.
- Synchronization can be safely turned off on border routers:
  - When all transit routers inside the AS are running fully meshed IBGP.
  - When the AS is not a transit AS.



Results:

- By itself, this will not change the reachability issues, because we have not resolved the problem with the boundary routers being able to reach the next-hop interface (SanJose1 still can not reach 172.30.1.1).
- We will make another change in a moment.
- Later, we might see that creating a full-mesh IBGP domain might not be the best configuration choice.

ISP1

```
router bgp 100
 network 10.2.0.0 mask 255.255.255.0
 neighbor 10.1.0.2 remote-as 200
```

ISP2

```
router bgp 300
 network 172.16.0.0
 neighbor 172.30.1.2 remote-as 200
```

SanJose3

```
router rip
 version 2
 network 192.168.1.0
!
router bgp 200
 no synchronization
 neighbor 192.168.1.193 remote-as 200
 neighbor 192.168.1.197 remote-as 200
```

SanJose1

```
router rip
 version 2
 network 192.168.1.0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 10.1.0.1 remote-as 100
 neighbor 192.168.1.194 remote-as 200
 neighbor 192.168.1.197 remote-as 200
!
```

SanJose2

```
router rip
 version 2
 network 192.168.1.0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 172.30.1.1 remote-as 300
 neighbor 192.168.1.193 remote-as 200
 neighbor 192.168.1.198 remote-as 200
```

Same Problem

SanJose1 still only sees the BGP routes distributed by itself and its peer in their BGP tables.

```
SanJose1#show ip route

10.0.0.0/24 is subnetted, 2 subnets
B       10.2.0.0 [20/0] via 10.1.0.1, 00:09:13
C       10.1.0.0 is directly connected, Serial0/1
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/2] via 192.168.1.194, 00:00:27, Serial0/0
C       192.168.1.64/27 is directly connected, Loopback0
R       192.168.1.32/27 [120/1] via 192.168.1.194, 00:00:27, Serial0/0
C       192.168.1.192/30 is directly connected, Serial0/0
R       192.168.1.196/30 [120/1] via 192.168.1.194, 00:00:27, Serial0/0

SanJose1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      10.1.0.1                 0             0 100 i
* i172.16.0.0       172.30.1.1               0    100      0 300 i
*> 192.168.1.0      0.0.0.0                  0         32768 i
* i                 192.168.1.197            0    100      0 i
```

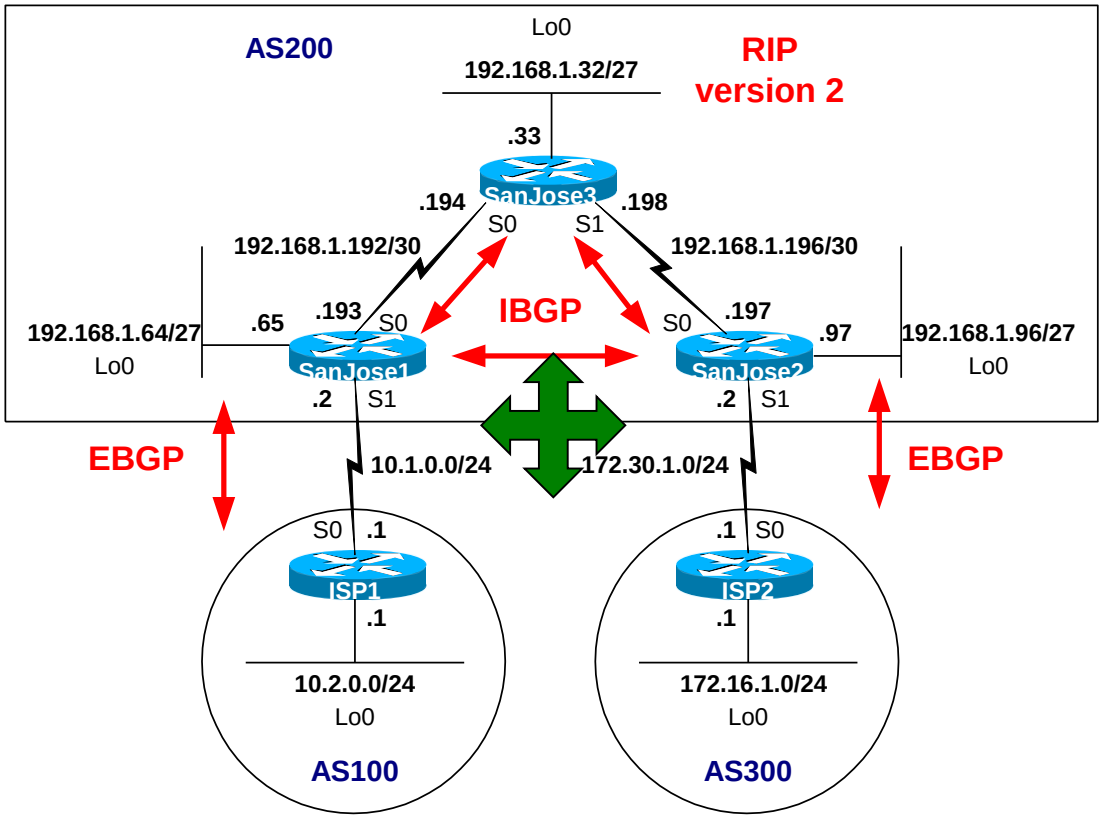
- **Even though 172.16.0.0 network is in BGP table it is not in routing table because there is no route to the next hop of 172.30.1.1**

Same Problem

SanJose2 still only sees the BGP routes distributed by itself and its peer in their BGP tables.

```
SanJose2#show ip route
B     172.16.0.0/16 [20/0] via 172.30.1.1, 00:11:27
      172.30.0.0/24 is subnetted, 1 subnets
C        172.30.1.0 is directly connected, Serial0/1
      192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
C        192.168.1.96/27 is directly connected, Loopback0
R        192.168.1.64/27 [120/2] via 192.168.1.198, 00:00:19, Serial0/0
R        192.168.1.32/27 [120/1] via 192.168.1.198, 00:00:19, Serial0/0
R        192.168.1.192/30 [120/1] via 192.168.1.198, 00:00:19, Serial0/0
C        192.168.1.196/30 is directly connected, Serial0/0

SanJose2#show ip bgp
Network          Next Hop            Metric LocPrf Weight Path
* i10.2.0.0/24     10.1.0.1                 0    100      0 100 i
*> 172.16.0.0      172.30.1.1               0             0 300 i
* i192.168.1.0     192.168.1.193            0    100      0 i
*>                 0.0.0.0                  0         32768 i
```

- **Even though 10.2.0.0 network is in BGP table it is not in routing table because there is no route to the next hop of 10.1.0.1**



Complete reachability

Step #4 – Adding the networks between the boundary routers and the ISP routers to the RIP routing domain.

- EBGP is configured between ISP routers and AS200 boundary routers.
- RIPv2 is running within AS200, **including the 10.0.0.0 and 172.30.0.0 network statements and passive interfaces**.
- IBGP is configured between SanJose1, SanJose2 and SanJose3, full-mesh.
- **No synchronization** has been configured on all IBGP routers.
- There are **no** static routes.

ISP1

```
router bgp 100
 network 10.2.0.0 mask 255.255.255.0
 neighbor 10.1.0.2 remote-as 200
```

ISP2

```
router bgp 300
 network 172.16.0.0
 neighbor 172.30.1.2 remote-as 200
```

SanJose3

```
router rip
 version 2
 network 192.168.1.0
!
router bgp 200
 no synchronization
 neighbor 192.168.1.193 remote-as 200
 neighbor 192.168.1.197 remote-as 200
```

SanJose1

```
router rip
 version 2
 passive-interface Serial0/1
 network 10.0.0.0
 network 192.168.1.0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 10.1.0.1 remote-as 100
 neighbor 192.168.1.194 remote-as 200
 neighbor 192.168.1.197 remote-as 200
```

SanJose2

```
router rip
 version 2
 passive-interface Serial0/1
 network 172.30.0.0
 network 192.168.1.0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 172.30.1.1 remote-as 300
 neighbor 192.168.1.193 remote-as 200
 neighbor 192.168.1.198 remote-as 200
```

```
SanJose1#show ip route
B     172.16.0.0/16 [200/0] via 172.30.1.1, 00:17:56
R     172.30.0.0/16 [120/2] via 192.168.1.194, 00:00:20, Serial0/0
      10.0.0.0/24 is subnetted, 2 subnets
B        10.2.0.0 [20/0] via 10.1.0.1, 00:17:50
C        10.1.0.0 is directly connected, Serial0/1
      192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R        192.168.1.96/27 [120/2] via 192.168.1.194, 00:00:20, Serial0/0
C        192.168.1.64/27 is directly connected, Loopback0
R        192.168.1.32/27 [120/1] via 192.168.1.194, 00:00:20, Serial0/0
C        192.168.1.192/30 is directly connected, Serial0/0
R        192.168.1.196/30 [120/1] via 192.168.1.194, 00:00:20, Serial0/0

SanJose1#show ip bgp
   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      10.1.0.1                 0             0 100 i
*>i172.16.0.0       172.30.1.1               0    100      0 300 i
*> 192.168.1.0      0.0.0.0                  0         32768 i
* i                  192.168.1.197            0    100      0 i
```

- **Now that the networks between the ISPs and the boundary routers are being propagated by RIP, the boundary routers now have a route to the BGP next hops.**
- **Because of this the BGP routes to the “opposite” ISP network is reachable and in the routing table.**

```
SanJose2#show ip route
B     172.16.0.0/16 [20/0] via 172.30.1.1, 00:21:51
      172.30.0.0/24 is subnetted, 1 subnets
C        172.30.1.0 is directly connected, Serial0/1
      10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
B        10.2.0.0/24 [200/0] via 10.1.0.1, 00:21:40
R        10.0.0.0/8 [120/2] via 192.168.1.198, 00:00:03, Serial0/0
      192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
C        192.168.1.96/27 is directly connected, Loopback0
R        192.168.1.64/27 [120/2] via 192.168.1.198, 00:00:03, Serial0/0
R        192.168.1.32/27 [120/1] via 192.168.1.198, 00:00:03, Serial0/0
R        192.168.1.192/30 [120/1] via 192.168.1.198, 00:00:03, Serial0/0
C        192.168.1.196/30 is directly connected, Serial0/0

SanJose2#show ip bgp
   Network          Next Hop            Metric LocPrf Weight Path
*>i10.2.0.0/24      10.1.0.1                 0    100      0 100 i
*> 172.16.0.0       172.30.1.1               0             0 300 i
* i192.168.1.0      192.168.1.193            0    100      0 i
*>                  0.0.0.0                  0         32768 i
```

- Not only is the 10.2.0.0 network reachable, but the 10.0.0.0/8 classful network, which includes 10.1.0.0/24. This will come into play in a moment.
- The BGP route to the "opposite" ISP network is reachable and in the routing table.

```
SanJose3#show ip route

B    172.16.0.0/16 [200/0] via 172.30.1.1, 00:20:06
R    172.30.0.0/16 [120/1] via 192.168.1.197, 00:00:12, Serial0/1
     10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
B       10.2.0.0/24 [200/0] via 10.1.0.1, 00:19:56
R       10.0.0.0/8 [120/1] via 192.168.1.193, 00:00:11, Serial0/0
     192.168.1.0/24 is variably subnetted, 6 subnets, 3 masks
R       192.168.1.96/27 [120/1] via 192.168.1.197, 00:00:12, Serial0/1
R       192.168.1.64/27 [120/1] via 192.168.1.193, 00:00:11, Serial0/0
C       192.168.1.32/27 is directly connected, Loopback0
B       192.168.1.0/24 [200/0] via 192.168.1.193, 00:19:53
C       192.168.1.192/30 is directly connected, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/1

SanJose3#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*>i10.2.0.0/24      10.1.0.1                 0    100      0 100 i
*>i172.16.0.0       172.30.1.1               0    100      0 300 i
*>i192.168.1.0      192.168.1.193            0    100      0 i
* i                 192.168.1.197            0    100      0 i
```

```
ISP1#show ip route

B     172.16.0.0/16 [20/0] via 10.1.0.2, 00:15:14
      10.0.0.0/24 is subnetted, 2 subnets
C        10.2.0.0 is directly connected, Loopback0
C        10.1.0.0 is directly connected, Serial0
B     192.168.1.0/24 [20/0] via 10.1.0.2, 00:15:14

ISP1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      0.0.0.0                  0         32768 i
*> 172.16.0.0       10.1.0.2                               0 200 300 i
*> 192.168.1.0      10.1.0.2                 0             0 200 i
```

- **The 172.16.0.0 network is now propagated from SanJose1 to ISP1.**

```
ISP2#show ip route

     172.16.0.0/24 is subnetted, 1 subnets
C       172.16.1.0 is directly connected, Loopback0
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/0
     10.0.0.0/24 is subnetted, 1 subnets
B       10.2.0.0 [20/0] via 172.30.1.2, 00:23:09
B    192.168.1.0/24 [20/0] via 172.30.1.2, 00:23:09

ISP2#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      172.30.1.2                             0 200 100 i
*> 172.16.0.0       0.0.0.0                  0         32768 i
*> 192.168.1.0      172.30.1.2               0             0 200 i
```

- **The 10.2.0.0 network (and the 10.0.0.0 network) is now propagated from SanJose1 to ISP1.**

```
SanJose1#ping 172.16.1.1
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 84/84/84 ms
SanJose1#ping 10.2.0.1
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 28/28/28 ms
-----------------------------------
SanJose3#ping 172.16.1.1
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 56/56/56 ms
SanJose3#ping 10.2.0.1
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 56/56/60 ms
-------------------------------------
SanJose2#ping 172.16.1.1
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 28/28/32 ms
SanJose2#ping 10.2.0.1
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 84/84/84 ms
```

```
ISP2#ping 10.2.0.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)
```

- **I thought you said all routers could reach all networks?**
- **What happened???**

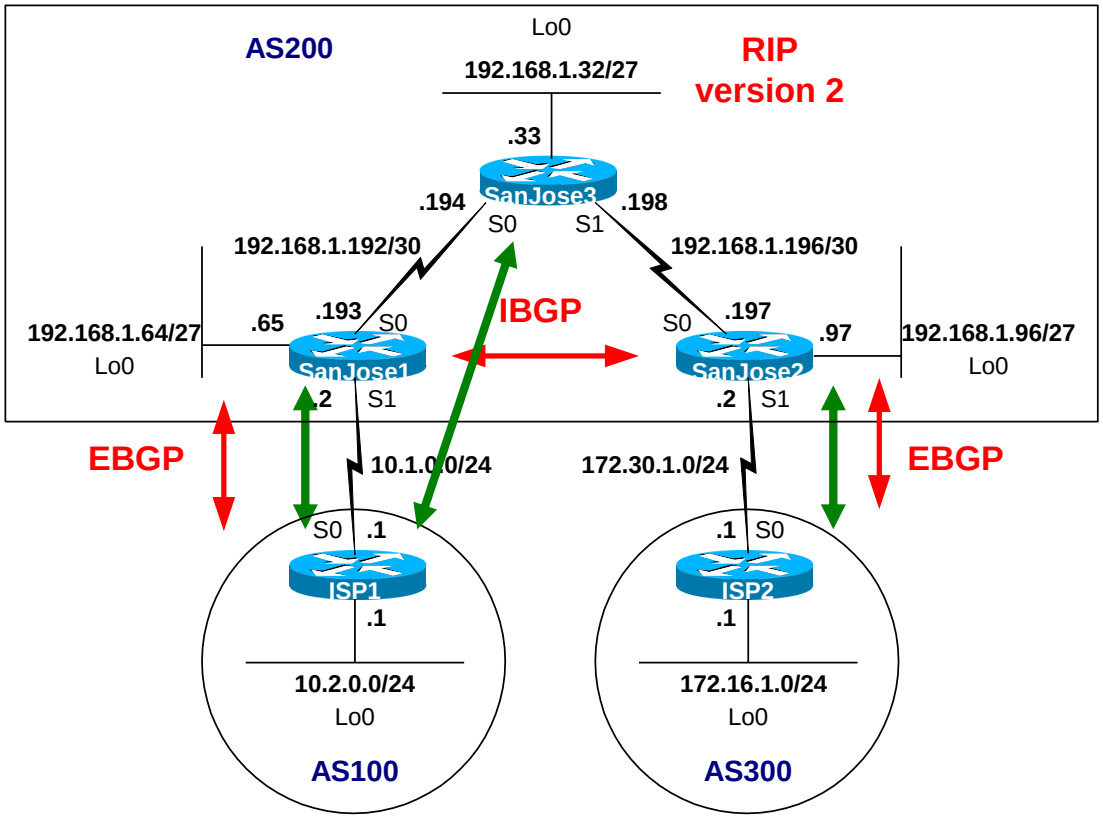
- **ISP2 can reach 10.2.0.1**
- **What do you think the problem is?**

```
ISP2#show ip route
     172.16.0.0/24 is subnetted, 1 subnets
C       172.16.1.0 is directly connected, Loopback0
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/0
     10.0.0.0/24 is subnetted, 1 subnets
B       10.2.0.0 [20/0] via 172.30.1.2, 00:23:09
B    192.168.1.0/24 [20/0] via 172.30.1.2, 00:23:09

ISP2#show ip bgp
   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      172.30.1.2                             0 200 100 i
*> 172.16.0.0       0.0.0.0                  0         32768 i
*> 192.168.1.0      172.30.1.2               0             0 200 i

ISP1#show ip route
B    172.16.0.0/16 [20/0] via 10.1.0.2, 00:15:14
     10.0.0.0/24 is subnetted, 2 subnets
C       10.2.0.0 is directly connected, Loopback0
C       10.1.0.0 is directly connected, Serial0
B    192.168.1.0/24 [20/0] via 10.1.0.2, 00:15:14

ISP1#show ip bgp
   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      0.0.0.0                  0         32768 i
*> 172.16.0.0       10.1.0.2                               0 200 300 i
*> 192.168.1.0      10.1.0.2                 0             0 200 I
```

- **ISP2 does have a route to 10.2.0.0 and ISP1 has a route to 172.16.0.0**
- **The ping uses a source IP address of 172.30.1.1 and ISP1 does NOT have a route back to this destination address. (This also applies to pings from ISP1.)**
- **Extended pings with a reachable source IP address will work!**

```
ISP2#ping
Protocol [ip]:
Target IP address: 10.2.0.1
Repeat count [5]:
Datagram size [100]:
Timeout in seconds [2]:
Extended commands [n]: y
Source address or interface: 172.16.1.1
Type of service [0]:
Set DF bit in IP header? [no]:
Validate reply data? [no]:
Data pattern [0xABCD]:
Loose, Strict, Record, Timestamp, Verbose[none]:
Sweep range of sizes [n]:
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 108/112/116 ms
```

- **ISP2 does have a route to 10.2.0.0 and ISP1 has a route to 172.16.0.0**
- **The ping uses a source IP address of 172.30.1.1 and ISP1 does NOT have a route back to this destination address. (This also applies to pings from ISP1.)**
- **Extended pings with a reachable source IP address will work!**



Scenario B: Step #1 – IBGP between SanJose1 and SanJose2 only

- EBGP is configured between ISP routers and AS200 boundary routers.
- RIPv2 is running within AS200, including the 10.0.0.0 and 172.30.0.0 network statements and passive interfaces.
- **IBGP is configured only between SanJose1, SanJose2.**
- No synchronization has been configured on all IBGP routers.
- There are **no** static routes.

ISP1

```
router bgp 100
 network 10.2.0.0 mask 255.255.255.0
 neighbor 10.1.0.2 remote-as 200
```

ISP2

```
router bgp 300
 network 172.16.0.0
 neighbor 172.30.1.2 remote-as 200
```

SanJose3

```
router rip
 version 2
 network 192.168.1.0
!
no router bgp 200
```

SanJose1

```
router rip
 version 2
 passive-interface Serial0/1
 network 10.0.0.0
 network 192.168.1.0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 10.1.0.1 remote-as 100
 neighbor 192.168.1.197 remote-as 200
```

SanJose2

```
router rip
 version 2
 passive-interface Serial0/1
 network 172.30.0.0
 network 192.168.1.0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 172.30.1.1 remote-as 300
 neighbor 192.168.1.193 remote-as 200
```

Results

- SanJose3 does **NOT** have a route to 172.16.0.0 on ISP2.
- SanJose3 does have a route to 10.0.0.0 on ISP1 only because SanJose1 is propagating its 10.1.0.0/24 subnet, so it has a route to the classful 10.0.0.0 network.
- SanJose3 does **NOT** have the specific ISP networks in its routing table so it may drop packets as it is passed through between SanJose1 and SanJose2.
  - SanJose1 and SanJose2 can **NOT** reach the opposite ISP2's 172.16.1.1 because SanJose3 does not have a route to 172.16.0.0 on ISP1.
  - However, SanJose1, SanJose2 and SanJose3 can reach ISP1's 10.2.0.1 because RIP is sending the 10.0.0.0 classful route which includes the serial link between SJ1 and ISP1.
- ISPs – They will **NOT** be able to ping each other, because SJ3 has a route for the 10.0.0.0 network but **NOT** for the 172.16.0.0 network, thus pings from ISP2 to ISP1 will get to ISP2 but the replies will get dropped by SJ3 as it does not have a route for 172.16.0.0.

```
ISP1#show ip route

B    172.16.0.0/16 [20/0] via 10.1.0.2, 00:34:13
     10.0.0.0/24 is subnetted, 2 subnets
C       10.2.0.0 is directly connected, Loopback0
C       10.1.0.0 is directly connected, Serial0
B    192.168.1.0/24 [20/0] via 10.1.0.2, 00:38:39

ISP1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      0.0.0.0                  0         32768 i
*> 172.16.0.0       10.1.0.2                               0 200 300 i
*> 192.168.1.0      10.1.0.2                 0             0 200 i


ISP1#ping 192.168.1.33
Sending 5, 100-byte ICMP Echos to 192.168.1.33, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 56/56/60 ms

ISP1#ping 172.16.1.1
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
U.U.U
Success rate is 0 percent (0/5)
```

This looks fine for ISP1, but lets look at the AS200 routers.

MAKE THIS AN EXTENDED PING

```
SanJose1#show ip route
B    172.16.0.0/16 [200/0] via 172.30.1.1, 00:35:40
R    172.30.0.0/16 [120/2] via 192.168.1.194, 00:00:04, Serial0/0
     10.0.0.0/24 is subnetted, 2 subnets
B       10.2.0.0 [20/0] via 10.1.0.1, 00:40:06
C       10.1.0.0 is directly connected, Serial0/1
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/2] via 192.168.1.194, 00:00:04, Serial0/0
C       192.168.1.64/27 is directly connected, Loopback0
R       192.168.1.32/27 [120/1] via 192.168.1.194, 00:00:04, Serial0/0
C       192.168.1.192/30 is directly connected, Serial0/0
R       192.168.1.196/30 [120/1] via 192.168.1.194, 00:00:05, Serial0/0

SanJose1#show ip bgp
Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      10.1.0.1                 0             0 100 i
*>i172.16.0.0       172.30.1.1               0    100      0 300 i
* i192.168.1.0      192.168.1.197            0    100      0 i
*>                  0.0.0.0                  0         32768 i

SanJose1#ping 10.2.0.1
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 28/28/32 ms

SanJose1#ping 172.16.1.1
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
U.U.U
Success rate is 0 percent (0/5)
```

This looks fine for SJ1, but lets look at SJ3.

```
SanJose3#show ip route
R    172.30.0.0/16 [120/1] via 192.168.1.197, 00:00:25, Serial0/1
R    10.0.0.0/8 [120/1] via 192.168.1.193, 00:00:06, Serial0/0
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/1] via 192.168.1.197, 00:00:25, Serial0/1
R       192.168.1.64/27 [120/1] via 192.168.1.193, 00:00:06, Serial0/0
C       192.168.1.32/27 is directly connected, Loopback0
C       192.168.1.192/30 is directly connected, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/1

SanJose3#ping 10.2.0.1
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 56/56/60 ms

SanJose3#ping 172.16.1.1
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)
```

Where's 172.16.0.0?

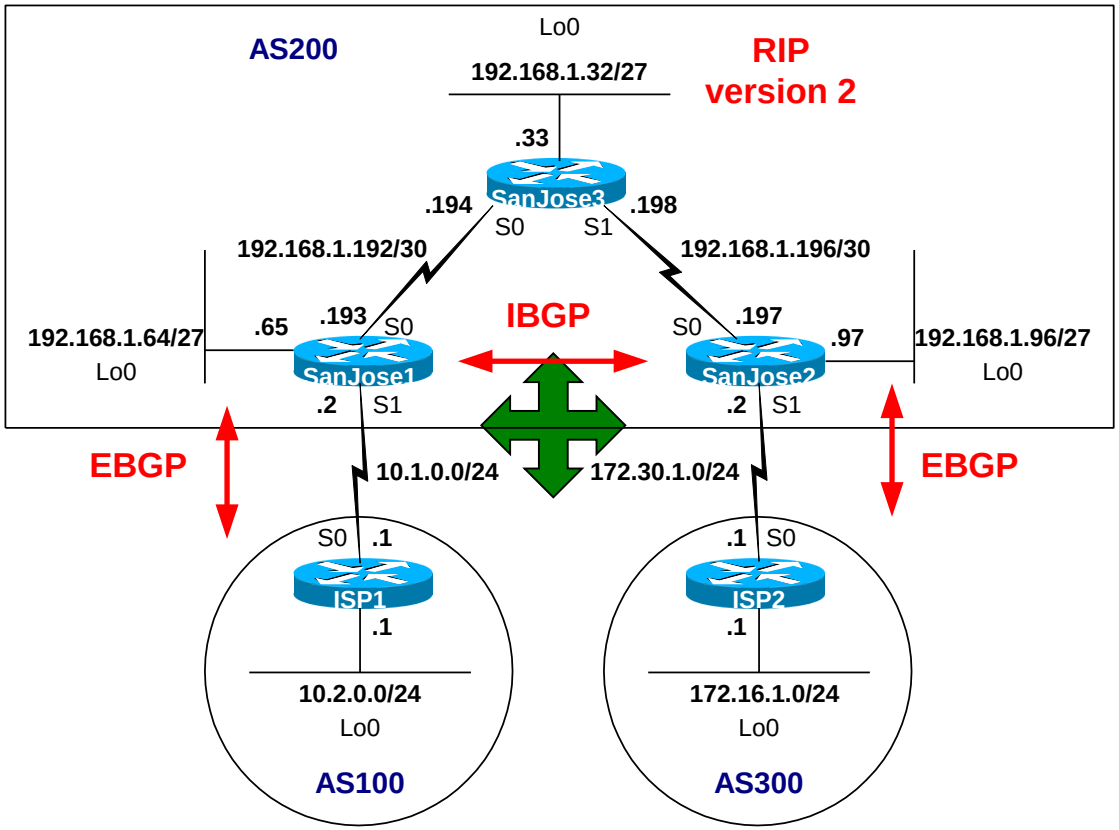
- **Notice that SanJose3 does NOT have a route for 172.16.0.0.**
- **When a packet with a destination address for this network reaches SanJose3 it drops it.**
- **Remember, we removed SanJose3 as an IBGP peer.**

```
SanJose2#show ip route
B    172.16.0.0/16 [20/0] via 172.30.1.1, 00:43:13
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/1
     10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
B       10.2.0.0/24 [200/0] via 10.1.0.1, 00:38:57
R       10.0.0.0/8 [120/2] via 192.168.1.198, 00:00:04, Serial0/0
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
C       192.168.1.96/27 is directly connected, Loopback0
R       192.168.1.64/27 [120/2] via 192.168.1.198, 00:00:04, Serial0/0
R       192.168.1.32/27 [120/1] via 192.168.1.198, 00:00:05, Serial0/0
R       192.168.1.192/30 [120/1] via 192.168.1.198, 00:00:05, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/0

SanJose2#show ip bgp
Network          Next Hop            Metric LocPrf Weight Path
*>i10.2.0.0/24     10.1.0.1                 0    100      0 100 i
*> 172.16.0.0      172.30.1.1               0             0 300 i
* i192.168.1.0     192.168.1.193            0    100      0 i
*>                 0.0.0.0                  0         32768 i

SanJose2#ping 10.2.0.1
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 84/84/84 ms

SanJose2#ping 172.16.1.1
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 28/29/32 ms
```

Again,this looks fine for SJ2, but 172.16.0.0 does NOT exist on SJ3.

```
ISP2#show ip route
172.16.0.0/24 is subnetted, 1 subnets
C       172.16.1.0 is directly connected, Loopback0
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/0
     10.0.0.0/24 is subnetted, 1 subnets
B       10.2.0.0 [20/0] via 172.30.1.2, 00:44:58
B    192.168.1.0/24 [20/0] via 172.30.1.2, 00:49:15

ISP2#show ip bgp
Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      172.30.1.2                             0 200 100 i
*> 172.16.0.0       0.0.0.0                  0         32768 i
*> 192.168.1.0      172.30.1.2               0             0 200 i

ISP2#ping 10.2.0.1
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)

ISP2#ping 192.168.1.65
Sending 5, 100-byte ICMP Echos to 192.168.1.65, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 84/84/88 ms
```

This also looks fine for ISP1, but remember the packets can't return to ISP2 because SJ3 does not have a route for 172.16.0.0.



Complete reachability

Step 2 - The Solution, Redistributing BGP routes into RIP

- EBGP is configured between ISP routers and AS200 boundary routers.
- RIPv2 is running within AS200, including the 10.0.0.0 and 172.30.0.0 network statements and passive interfaces.
- **IBGP is configured only between SanJose1, SanJose2.**
- No synchronization has been configured on all IBGP routers.
- **172.16.0.0 BGP routes are redistributed into AS200 via RIP**
- There are **no** static routes.

ISP1

```
router bgp 100
 network 10.2.0.0 mask 255.255.255.0
 neighbor 10.1.0.2 remote-as 200
```

ISP2

```
router bgp 300
 network 172.16.0.0
 neighbor 172.30.1.2 remote-as 200
```

SanJose3

```
router rip
 version 2
 network 192.168.1.0
```

SanJose1

```
router rip
 version 2
 passive-interface Serial0/1
 network 10.0.0.0
 network 192.168.1.0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 10.1.0.1 remote-as 100
 neighbor 192.168.1.197 remote-as 200
```

SanJose2

```
router rip
 version 2
 redistribute bgp 200 metric 2
 passive-interface Serial0/1
 network 172.30.0.0
 network 192.168.1.0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 172.30.1.1 remote-as 300
 neighbor 192.168.1.193 remote-as 200
!
```

```
SanJose2#show ip route
B     172.16.0.0/16 [20/0] via 172.30.1.1, 01:22:20
      172.30.0.0/24 is subnetted, 1 subnets
C        172.30.1.0 is directly connected, Serial0/1
      10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
B        10.2.0.0/24 [200/0] via 10.1.0.1, 01:18:04
R        10.0.0.0/8 [120/2] via 192.168.1.198, 00:00:21, Serial0/0
      192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
C        192.168.1.96/27 is directly connected, Loopback0
R        192.168.1.64/27 [120/2] via 192.168.1.198, 00:00:21, Serial0/0
R        192.168.1.32/27 [120/1] via 192.168.1.198, 00:00:21, Serial0/0
R        192.168.1.192/30 [120/1] via 192.168.1.198, 00:00:21, Serial0/0
C        192.168.1.196/30 is directly connected, Serial0/0

SanJose2#show ip bgp
BGP table version is 4, local router ID is 192.168.1.97
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*>i10.2.0.0/24      10.1.0.1                 0    100      0 100 i
*> 172.16.0.0       172.30.1.1               0             0 300 i
* i192.168.1.0      192.168.1.193            0    100      0 i
*>                  0.0.0.0                  0         32768 i
SanJose2#
```

As before.

```
ISP2#show ip route
172.16.0.0/24 is subnetted, 1 subnets
C       172.16.1.0 is directly connected, Loopback0
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/0
     10.0.0.0/24 is subnetted, 1 subnets
B       10.2.0.0 [20/0] via 172.30.1.2, 01:19:44
B    192.168.1.0/24 [20/0] via 172.30.1.2, 01:24:01

ISP2#show ip bgp
Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      172.30.1.2                             0 200 100 i
*> 172.16.0.0       0.0.0.0                  0         32768 i
*> 192.168.1.0      172.30.1.2               0             0 200 i
```

As before.

```
SanJose3#show ip route

R    172.16.0.0/16 [120/2] via 192.168.1.197, 00:00:04, Serial0/1
R    172.30.0.0/16 [120/1] via 192.168.1.197, 00:00:04, Serial0/1
R    10.0.0.0/8 [120/1] via 192.168.1.193, 00:00:25, Serial0/0
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/1] via 192.168.1.197, 00:00:04, Serial0/1
R       192.168.1.64/27 [120/1] via 192.168.1.193, 00:00:25, Serial0/0
C       192.168.1.32/27 is directly connected, Loopback0
C       192.168.1.192/30 is directly connected, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/1
SanJose3#ping 172.16.1.1
```

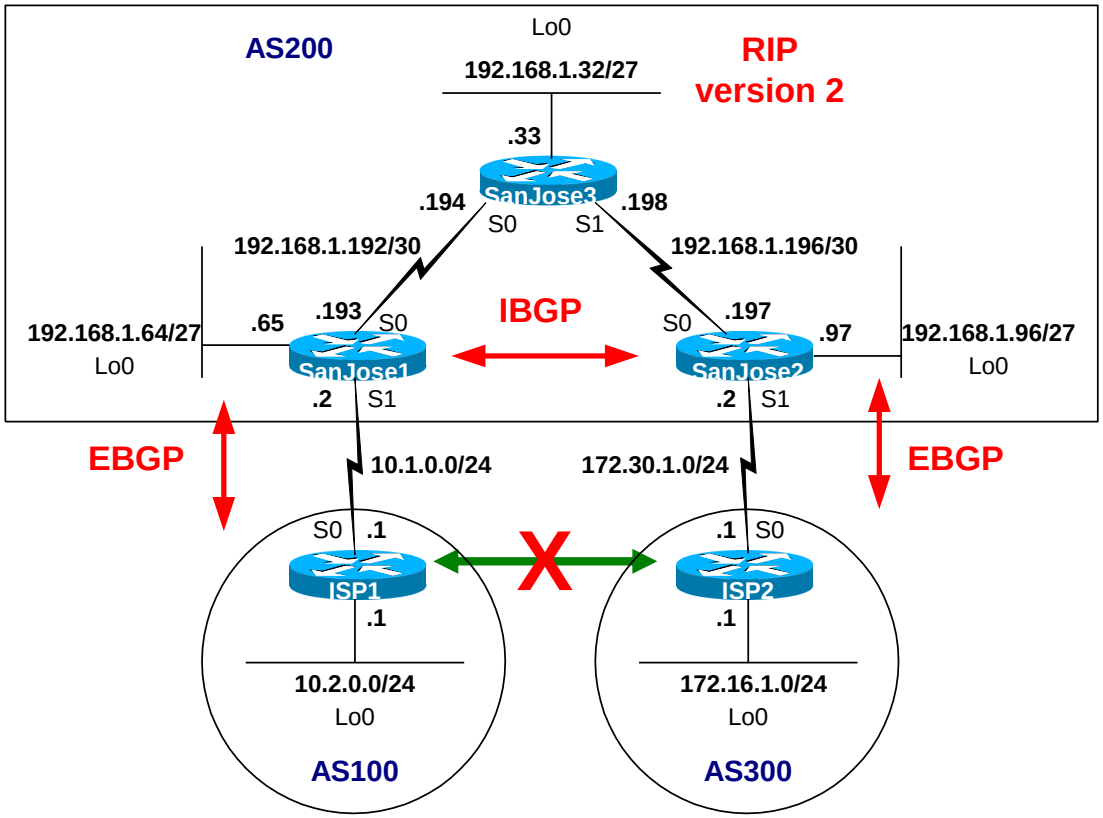
- SanJose3 now has a route for **172.16.0.0** which was redistributed into RIP by SanJose2.
  - This gives us the rest of the connectivity we need.
- 10.0.0.0 network was already in SanJose2's routing table because SanJose1's has a RIP network 10.0.0.0 statement.
  - Remember 10.2.0.0 and 10.1.0.0 belong to this classful network.

```
SanJose1#show ip route

R    172.16.0.0/16 [120/3] via 192.168.1.194, 00:00:02, Serial0/0
R    172.30.0.0/16 [120/2] via 192.168.1.194, 00:00:02, Serial0/0
     10.0.0.0/24 is subnetted, 2 subnets
B       10.2.0.0 [20/0] via 10.1.0.1, 01:28:18
C       10.1.0.0 is directly connected, Serial0/1
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/2] via 192.168.1.194, 00:00:02, Serial0/0
C       192.168.1.64/27 is directly connected, Loopback0
R       192.168.1.32/27 [120/1] via 192.168.1.194, 00:00:02, Serial0/0
C       192.168.1.192/30 is directly connected, Serial0/0
R       192.168.1.196/30 [120/1] via 192.168.1.194, 00:00:02, Serial0/0
SanJose1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      10.1.0.1                 0             0 100 i
*>i172.16.0.0       172.30.1.1               0    100      0 300 i
* i192.168.1.0      192.168.1.197            0    100      0 i
*>                  0.0.0.0                  0         32768 i
```

- **172.16.0.0** is now redistributed via RIP which has a lower administrative distance (120) than IBGP (200); (EBGP is 20), which was how this route was in SanJose1's routing table prior to redistribution by SanJose2..

```
ISP1#show ip route

B    172.16.0.0/16 [20/0] via 10.1.0.2, 01:25:26
     10.0.0.0/24 is subnetted, 2 subnets
C       10.2.0.0 is directly connected, Loopback0
C       10.1.0.0 is directly connected, Serial0
B    192.168.1.0/24 [20/0] via 10.1.0.2, 01:29:53

ISP1#show ip bgp

Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      0.0.0.0                  0         32768 i
*> 172.16.0.0       10.1.0.2                               0 200 300 i
*> 192.168.1.0      10.1.0.2                 0             0 200 i
```

As before.



Scenario C: Another Way – Step #1

- What if we did not want to redistribute the serial networks between the boundary routers and the ISPs into our rip network. What would be the consequences? How could we fix them?
- Eventually, for ISP1 to reach ISP2 we will use next-hop attribute.
- SanJose1 can ping 172.16.1.1. We will see that SanJose1 has an IGP route to 172.16.0.0 as this was redistributed by SanJose2 from BGP into rip, so it can ping 172.16.1.1. However, the 172.16.0.0 route in the BGP routing table shows a next hop of 172.30.1.1, in which SanJose1 does not have a route for. Because BGP does not have a route for the next hop, it does not include this route in its update to its EBGP peer, ISP1.
- ISP2 does have a route to ISP1's 10.2.0.0/24 network because RIP is advertising the classful 10.0.0.0 network which includes the serial link between ISP1 and SanJose1. However, it can NOT ping it because the reply can NOT get back as explained in the previous paragraph.

ISP1

```
router bgp 100
 network 10.2.0.0 mask 255.255.255.0
 neighbor 10.1.0.2 remote-as 200
```

ISP2

```
router bgp 300
 network 172.16.0.0
 neighbor 172.30.1.2 remote-as 200
```

SanJose3

```
router rip
 version 2
 network 192.168.1.0
```

SanJose1

```
router rip
   version 2
   redistribute bgp 200 metric 2
   passive-interface Serial0/1
   no network 10.0.0.0
   network 192.168.1.0
router bgp 200
   no synchronization
   network 192.168.1.0
   neighbor 10.1.0.1 remote-as 100
   neighbor 192.168.1.197 remote-as 200
```

SanJose2

```
router rip
   version 2
   redistribute bgp 200 metric 2
   passive-interface Serial0/1
   no network 172.30.0.0
   network 192.168.1.0
router bgp 200
   no synchronization
   network 192.168.1.0
   neighbor 172.30.1.1 remote-as 300
   neighbor 192.168.1.193 remote-as 200
!
```

```
ISP1#show ip route

     10.0.0.0/24 is subnetted, 2 subnets
C       10.2.0.0 is directly connected, Loopback0
C       10.1.0.0 is directly connected, Serial0
B    192.168.1.0/24 [20/0] via 10.1.0.2, 00:08:34

ISP1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      0.0.0.0                  0         32768 i
*> 192.168.1.0      10.1.0.2                 0             0 200 i

ISP1#ping 192.168.1.97
Sending 5, 100-byte ICMP Echos to 192.168.1.97, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 84/84/84 ms
ISP1#ping
Protocol [ip]:
Target IP address: 172.16.1.1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)

ISP1#ping 172.30.1.1  (Note: This is normal behavior and may not be necessary.)

Sending 5, 100-byte ICMP Echos to 172.30.1.1, timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)
```

Notice: No route for 172.16.0.0/16 (we'll see why)

```
SanJose1#show ip route

R    172.16.0.0/16 [120/3] via 192.168.1.194, 00:00:19, Serial0/0
     10.0.0.0/24 is subnetted, 2 subnets
B       10.2.0.0 [20/0] via 10.1.0.1, 01:20:45
C       10.1.0.0 is directly connected, Serial0/1
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/2] via 192.168.1.194, 00:00:19, Serial0/0
C       192.168.1.64/27 is directly connected, Loopback0
R       192.168.1.32/27 [120/1] via 192.168.1.194, 00:00:19, Serial0/0
C       192.168.1.192/30 is directly connected, Serial0/0
R       192.168.1.196/30 [120/1] via 192.168.1.194, 00:00:19, Serial0/0
SanJose1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      10.1.0.1                 0             0 100 i
* i172.16.0.0       172.30.1.1               0    100      0 300 i
*> 192.168.1.0      0.0.0.0                  0         32768 i
* i                 192.168.1.197            0    100      0 i
SanJose1#ping 10.2.0.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 28/29/32 ms
SanJose1#ping 172.16.1.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 84/84/84 ms
```

Known via RIP because is was redistributed from BGP.

NOT sent to ISP1 via EBGP because Next hop is not reachable.

EBGP routes carried into IBGP use the external EBGP next hop.

```
SanJose3#show ip route

R    172.16.0.0/16 [120/2] via 192.168.1.197, 00:00:09, Serial0/1
R    10.0.0.0/8 [120/1] via 192.168.1.193, 00:00:22, Serial0/0
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/1] via 192.168.1.197, 00:00:09, Serial0/1
R       192.168.1.64/27 [120/1] via 192.168.1.193, 00:00:22, Serial0/0
C       192.168.1.32/27 is directly connected, Loopback0
C       192.168.1.192/30 is directly connected, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/1
SanJose3#

SanJose3#ping 10.2.0.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 56/57/64 ms
SanJose3#ping 172.16.1.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 56/56/56 ms
```

```
SanJose2#show ip route

B    172.16.0.0/16 [20/0] via 172.30.1.1, 01:23:21
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/1
     10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
B       10.2.0.0/24 [200/0] via 10.1.0.1, 00:31:24
R       10.0.0.0/8 [120/2] via 192.168.1.198, 00:00:25, Serial0/0
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
C       192.168.1.96/27 is directly connected, Loopback0
R       192.168.1.64/27 [120/2] via 192.168.1.198, 00:00:25, Serial0/0
R       192.168.1.32/27 [120/1] via 192.168.1.198, 00:00:25, Serial0/0
R       192.168.1.192/30 [120/1] via 192.168.1.198, 00:00:25, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/0

SanJose2#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*>i10.2.0.0/24      10.1.0.1                 0    100      0 100 i
*> 172.16.0.0       172.30.1.1               0             0 300 i
* i192.168.1.0      192.168.1.193            0    100      0 i
*>                  0.0.0.0                  0         32768 i
SanJose2#ping 10.2.0.1
Sending 5, 100-byte ICMP Echos to 10.2.0.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 84/84/84 ms
SanJose2#ping 172.16.1.1

Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 28/29/32 ms
```

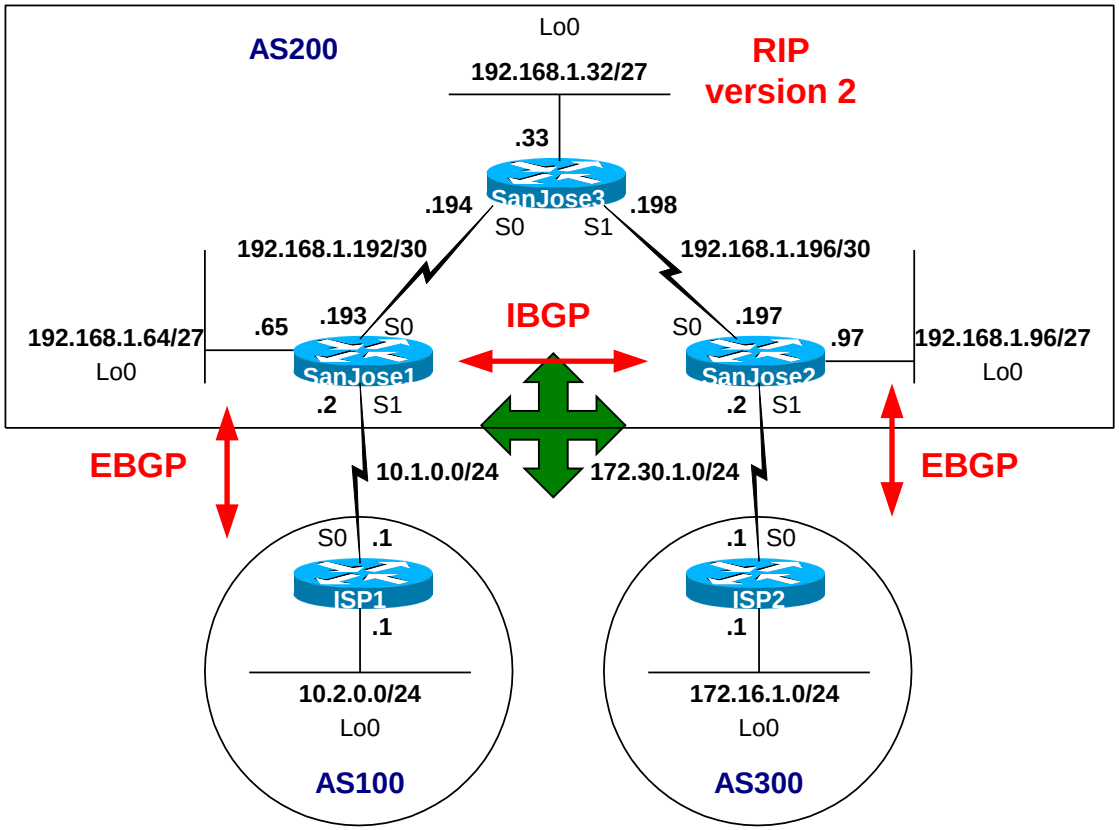
SanJose2 sends 10.2.0.0 to ISP2 via EBGP because Next hop of 10.1.0.1 is reachable, as part of the 10.0.0.0/8 net.

RIP redistributed the entire classful 10.0.0.0/8 network.

```
ISP2#show ip route

     172.16.0.0/24 is subnetted, 1 subnets
C       172.16.1.0 is directly connected, Loopback0
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/0
     10.0.0.0/24 is subnetted, 1 subnets
B       10.2.0.0 [20/0] via 172.30.1.2, 00:33:01
B    192.168.1.0/24 [20/0] via 172.30.1.2, 01:25:53
```

Notice: There is a route for 10.2.0.0

```
ISP2#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      172.30.1.2                             0 200 100 i
*> 172.16.0.0       0.0.0.0                  0         32768 i
*> 192.168.1.0      172.30.1.2               0             0 200 i
ISP2#ping
Protocol [ip]:
Target IP address: 10.2.0.1
.....
Success rate is 0 percent (0/5)
```



Complete reachability

Step #2 - The Solution – next-hop-self

- EBGP routes carried into IBGP use the external EBGP next hop. SanJose1's BGP table has 172.30.1.1 as the next hop for 172.16.0.0/16, but does not have a route to 172.30.1.1, so it does not forward this route to ISP1.
- BGP Path Selection: "If the path specifies a next hop that is inaccessible, drop the update." (coming)
- Lets change the next hop to be the internal IBGP next hop.
- IBGP peers SanJose1 and SanJose2 will use next-hop self in peering with each other. This will change to next hops from the EBGP next hop to the internal IBGP next hop.
- All works!!!

ISP1

```
router bgp 100
 network 10.2.0.0 mask 255.255.255.0
 neighbor 10.1.0.2 remote-as 200
```

ISP2

```
router bgp 300
 network 172.16.0.0
 neighbor 172.30.1.2 remote-as 200
```

SanJose3

```
router rip
 version 2
 network 192.168.1.0
```

SanJose1

```
router rip
 version 2
 redistribute bgp 200 metric 2
 passive-interface Serial0/1
 no network 10.0.0.0 (from before)
 network 192.168.1.0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 10.1.0.1 remote-as 100
 neighbor 192.168.1.197 remote-as 200
 neighbor 192.168.1.197 next-hop-self
```

SanJose2

```
router rip
 version 2
 redistribute bgp 200 metric 2
 passive-interface Serial0/1
 no network 172.30.0.0 (from before)
 network 192.168.1.0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 172.30.1.1 remote-as 300
 neighbor 192.168.1.193 remote-as 200
 neighbor 192.168.1.193 next-hop-self
!
```

```
ISP1#show ip route

B    172.16.0.0/16 [20/0] via 10.1.0.2, 00:03:49
     10.0.0.0/24 is subnetted, 2 subnets
C       10.2.0.0 is directly connected, Loopback0
C       10.1.0.0 is directly connected, Serial0
B    192.168.1.0/24 [20/0] via 10.1.0.2, 00:03:22

ISP1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      0.0.0.0                  0         32768 i
*> 172.16.0.0       10.1.0.2                               0 200 300 i
*> 192.168.1.0      10.1.0.2                 0             0 200 i

ISP1#ping 192.168.1.33
Sending 5, 100-byte ICMP Echos to 192.168.1.33, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 56/56/60 ms

ISP1#ping
Protocol [ip]:
Target IP address: 172.16.1.1
Extended commands [n]: y
Source address or interface: 10.2.0.1
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 108/111/112 ms
ISP1#
```

Looks good!

SanJose2: neighbor 192.168.1.193 next-hop-self (->192.168.1.197)

```
SanJose1#show ip route

R    172.16.0.0/16 [120/3] via 192.168.1.194, 00:00:24, Serial0/0
     10.0.0.0/24 is subnetted, 2 subnets
B       10.2.0.0 [20/0] via 10.1.0.1, 00:05:51
C       10.1.0.0 is directly connected, Serial0/1
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/2] via 192.168.1.194, 00:00:24, Serial0/0
C       192.168.1.64/27 is directly connected, Loopback0
R       192.168.1.32/27 [120/1] via 192.168.1.194, 00:00:24, Serial0/0
R       192.168.1.32/27 [120/1] via 192.168.1.194, 00:00:24, Serial0/0
R       192.168.1.196/30 [120/1] via 192.168.1.194, 00:00:25, Serial0/0
```

Looks good!

The route to the next hop

```
SanJose1#show ip bgp
```

Effect of the next-hop-self (was 172.30.1.1)

```
   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      10.1.0.1                 0             0 100 i
*>i172.16.0.0       192.168.1.197            0    100      0 300 i
* i192.168.1.0      192.168.1.197            0    100      0 i
*>                  0.0.0.0                  0         32768 i
```

```
SanJose3#show ip route

R    172.16.0.0/16 [120/2] via 192.168.1.197, 00:00:05, Serial0/1
R    10.0.0.0/8 [120/1] via 192.168.1.193, 00:00:26, Serial0/0
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
R       192.168.1.96/27 [120/1] via 192.168.1.197, 00:00:05, Serial0/1
R       192.168.1.64/27 [120/1] via 192.168.1.193, 00:00:27, Serial0/0
C       192.168.1.32/27 is directly connected, Loopback0
C       192.168.1.192/30 is directly connected, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/1
```

SanJose1: neighbor 192.168.1.197 next-hop-self (->192.168.1.193)

```
SanJose2#show ip route

B    172.16.0.0/16 [20/0] via 172.30.1.1, 00:09:30
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/1
     10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
B       10.2.0.0/24 [200/0] via 192.168.1.193, 00:09:23
R       10.0.0.0/8 [120/2] via 192.168.1.198, 00:00:17, Serial0/0
     192.168.1.0/24 is variably subnetted, 5 subnets, 2 masks
C       192.168.1.96/27 is directly connected, Loopback0
R       192.168.1.64/27 [120/2] via 192.168.1.198, 00:00:17, Serial0/0
R       192.168.1.32/27 [120/1] via 192.168.1.198, 00:00:17, Serial0/0
R       192.168.1.192/30 [120/1] via 192.168.1.198, 00:00:17, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/0
```

The route to the next hop

Affect of the next-hop-self (was 10.1.0.1)

```
SanJose2#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*>i10.2.0.0/24      192.168.1.193            0    100      0 100 i
*> 172.16.0.0       172.30.1.1               0             0 300 i
* i192.168.1.0      192.168.1.193            0    100      0 i
*>                  0.0.0.0                  0         32768 i
```

More to come…

- In the near future I will expand this example network to include many more examples including the use of attributes.
- For now:
  - What if we used a 172.16.1.0/24 mask on ISP2?
  - What if we were running OSPF instead of RIP.
  - More later…

What if we used a 172.16.1.0/24 mask on ISP2?

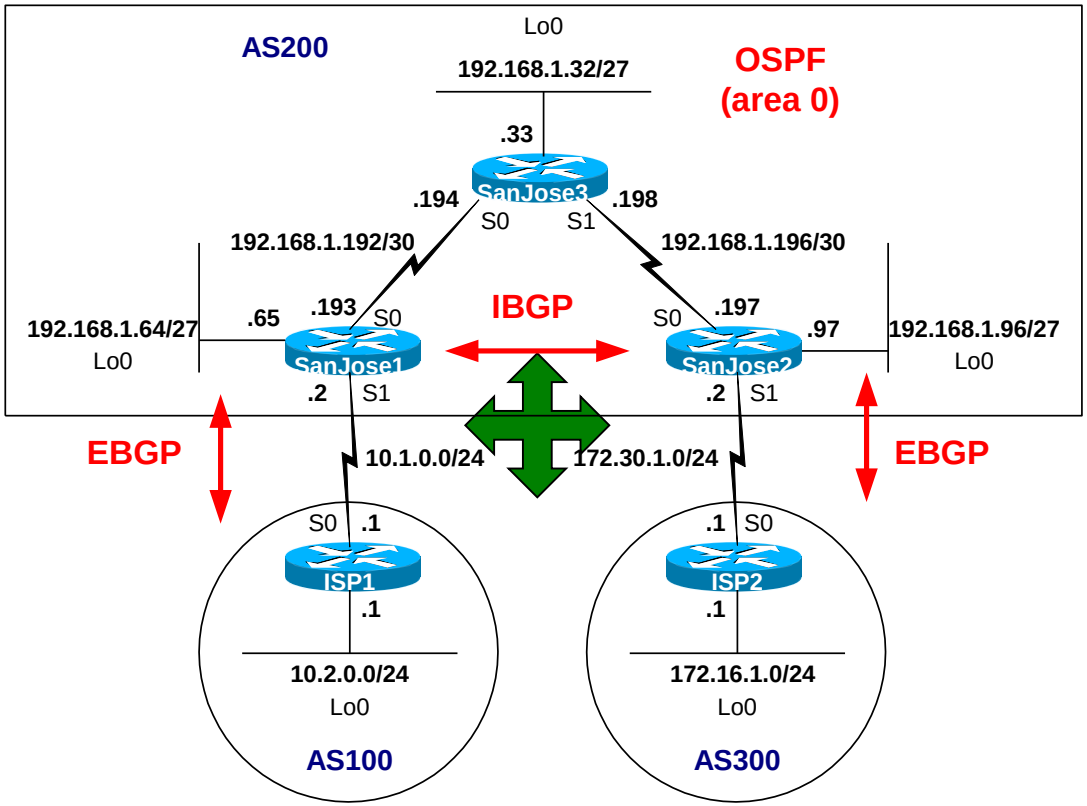
- 172.16.0.0/16 route distributed via RIP and BGP.

```
ISP2  BEFORE
router bgp 300
 network 172.16.0.0
 neighbor 172.30.1.2 remote-as 200

SanJose2#show ip route
B    172.16.0.0/16 [20/0] via 172.30.1.1, 00:09:30

SanJose3#show ip route
R    172.16.0.0/16 [120/2] via 192.168.1.197, 00:00:05, Serial0/1

SanJose1#show ip route
R    172.16.0.0/16 [120/3] via 192.168.1.194, 00:00:24, Serial0/0

ISP1#show ip route
B    172.16.0.0/16 [20/0] via 10.1.0.2, 00:03:49
```

What if we used a 172.16.1.0/24 mask on ISP2?

- 172.16.0.0/16 route distributed via RIP and 172.16.1.0/24 distributed via BGP.

```
ISP2  AFTER
router bgp 300
 network 172.16.1.0 mask 255.255.255.0
 neighbor 172.30.1.2 remote-as 200

SanJose2#show ip route
B    172.16.1.0/24 [20/0] via 172.30.1.1, 00:00:35

SanJose3#show ip route
R    172.16.0.0/16 [120/2] via 192.168.1.197, 00:00:05, Serial0/1

SanJose1#show ip route
R    172.16.0.0/16 [120/3] via 192.168.1.194, 00:00:24, Serial0/0
B    172.16.1.0/24 [20/0] via 192.168.1.197, 00:00:32

ISP1#show ip route
B    172.16.1.0/16 [20/0] via 10.1.0.2, 00:02:49
```



Complete reachability

Let's see how OSPF would work with BGP…

- **Notice that the 172.16.1.0/24 network is distributed within OSPF and not the classful 172.16.0.0/16 like with RIP.**
- **All works!!!**

ISP1

```
router bgp 100
 network 10.2.0.0 mask 255.255.255.0
 neighbor 10.1.0.2 remote-as 200
```

ISP2

```
router bgp 300
 network 172.16.0.0 mask 255.255.255.0
 neighbor 172.30.1.2 remote-as 200
```

SanJose3

```
router ospf 1
 network 192.168.1.33 0.0.0.0 area 0
 network 192.168.1.194 0.0.0.0 area 0
 network 192.168.1.198 0.0.0.0 area 0
```

SanJose1

```
router ospf 1
 redistribute bgp 200 metric 20 subnets
 network 192.168.1.65 0.0.0.0 area 0
 network 192.168.1.193 0.0.0.0 area 0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 10.1.0.1 remote-as 100
 neighbor 192.168.1.197 remote-as 200
 neighbor 192.168.1.197 next-hop-self
!
access-list 1 permit 10.2.0.0 0.0.0.255
```

SanJose2

```
router ospf 1
 redistribute bgp 200 metric 20 subnets
 network 192.168.1.97 0.0.0.0 area 0
 network 192.168.1.197 0.0.0.0 area 0
!
router bgp 200
 no synchronization
 network 192.168.1.0
 neighbor 172.30.1.1 remote-as 300
 neighbor 192.168.1.193 remote-as 200
 neighbor 192.168.1.193 next-hop-self
```

```
ISP2#show ip route

     172.16.0.0/24 is subnetted, 1 subnets
C       172.16.1.0 is directly connected, Loopback0
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/0
     10.0.0.0/24 is subnetted, 1 subnets
B       10.2.0.0 [20/0] via 172.30.1.2, 00:26:24
B    192.168.1.0/24 [20/0] via 172.30.1.2, 00:26:24

ISP2#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      172.30.1.2                             0 200 100 i
*> 172.16.1.0/24    0.0.0.0                  0         32768 i
*> 192.168.1.0      172.30.1.2               0             0 200 i
```

```
SanJose2#show ip route

     172.16.0.0/24 is subnetted, 1 subnets
B       172.16.1.0 [20/0] via 172.30.1.1, 00:25:24
     172.30.0.0/24 is subnetted, 1 subnets
C       172.30.1.0 is directly connected, Serial0/1
     10.0.0.0/24 is subnetted, 1 subnets
O E2    10.2.0.0 [110/20] via 192.168.1.198, 00:04:02, Serial0/0
     192.168.1.0/24 is variably subnetted, 5 subnets, 3 masks
C       192.168.1.96/27 is directly connected, Loopback0
O       192.168.1.65/32 [110/846] via 192.168.1.198, 00:09:46, Serial0/0
O       192.168.1.33/32 [110/782] via 192.168.1.198, 00:09:47, Serial0/0
O       192.168.1.192/30 [110/845] via 192.168.1.198, 00:09:47, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/0

SanJose2#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*>i10.2.0.0/24      192.168.1.193            0    100      0 100 i
*> 172.16.1.0/24    172.30.1.1               0             0 300 i
*> 192.168.1.0      0.0.0.0                  0         32768 i
* i                 192.168.1.193            0    100      0 i
```

```
SanJose3#show ip route

     172.16.0.0/24 is subnetted, 1 subnets
O E2    172.16.1.0 [110/20] via 192.168.1.197, 00:01:53, Serial0/1
     10.0.0.0/24 is subnetted, 1 subnets
O E2    10.2.0.0 [110/20] via 192.168.1.193, 00:02:23, Serial0/0
     192.168.1.0/24 is variably subnetted, 5 subnets, 3 masks
O       192.168.1.97/32 [110/65] via 192.168.1.197, 00:11:01, Serial0/1
O       192.168.1.65/32 [110/65] via 192.168.1.193, 00:11:01, Serial0/0
C       192.168.1.32/27 is directly connected, Loopback0
C       192.168.1.192/30 is directly connected, Serial0/0
C       192.168.1.196/30 is directly connected, Serial0/1
```

```
SanJose1#show ip route

     172.16.0.0/24 is subnetted, 1 subnets
O E2    172.16.1.0 [110/20] via 192.168.1.194, 00:02:08, Serial0/0
     10.0.0.0/24 is subnetted, 2 subnets
B       10.2.0.0 [20/0] via 10.1.0.1, 00:28:02
C       10.1.0.0 is directly connected, Serial0/1
     192.168.1.0/24 is variably subnetted, 5 subnets, 3 masks
O       192.168.1.97/32 [110/846] via 192.168.1.194, 00:11:09, Serial0/0
C       192.168.1.64/27 is directly connected, Loopback0
O       192.168.1.33/32 [110/782] via 192.168.1.194, 00:11:09, Serial0/0
C       192.168.1.192/30 is directly connected, Serial0/0
O       192.168.1.196/30 [110/845] via 192.168.1.194, 00:11:09, Serial0/0

SanJose1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      10.1.0.1                 0             0 100 i
*>i172.16.1.0/24    192.168.1.197            0    100      0 300 i
* i192.168.1.0      192.168.1.197            0    100      0 i
*>                  0.0.0.0                  0         32768 i
```

```
SanJose1#show ip ospf data

       OSPF Router with ID (192.168.1.65) (Process ID 1)


                Router Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum Link count
192.168.1.33    192.168.1.33    845         0x80000005 0xAB78   5
192.168.1.65    192.168.1.65    921         0x80000004 0x8883   3
192.168.1.97    192.168.1.97    801         0x80000004 0xB8EA   3

                Type-5 AS External Link States

Link ID         ADV Router      Age         Seq#       Checksum Tag
10.2.0.0        192.168.1.65    193         0x80000001 0xC1C2   100
172.16.1.0      192.168.1.97    165         0x80000001 0x28C1   300
```

```
ISP1#show ip route

     172.16.0.0/24 is subnetted, 1 subnets
B       172.16.1.0 [20/0] via 10.1.0.2, 00:25:00
     10.0.0.0/24 is subnetted, 2 subnets
C       10.2.0.0 is directly connected, Loopback0
C       10.1.0.0 is directly connected, Serial0
B    192.168.1.0/24 [20/0] via 10.1.0.2, 00:29:04

ISP1#show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.2.0.0/24      0.0.0.0                  0         32768 i
*> 172.16.1.0/24    10.1.0.2                               0 200 300 i
*> 192.168.1.0      10.1.0.2                 0             0 200 i
```

To be continued…